Three Essays in Economics of Climate Change

Dissertation
submitted to the
Faculty of Business, Economics and Informatics
of the University of Zurich

to obtain the degree of
Doktorin der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

Alena Miftakhova
from Russia

approved in July 2019 at the request of
Prof. Karl Schmedders, PhD
Prof. Ole Wilms, PhD

The Faculty of Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 17.07.2019

Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Acknowledgments

I express my deepest gratitude to my supervisor Karl Schmedders and to Kenneth Judd and Thomas Lontzek for their devoted teaching, guidance, and lasting support on this fascinating and challenging journey. I am also very grateful to Walter Pohl for our cooperation and to Dave Brooks for his invaluable editorial support. I am thankful to my colleagues Vanessa Kummer and Robert Erbe for our lively discussions, which enriched my life at our chair. I sincerely thank my family for their sympathy and continuous encouragement and Abi for being there for me.

This work is financially supported by the Swiss National Science Foundation.

Contents

Part I

Introduction

General Introduction

Human activities have been proven to have a strong impact on the climate system; the induced changes to the climate, in turn, have hazardous feedbacks on human welfare. Economics of climate change addresses the challenging tasks of estimating the costs associated with the human-induced changes to the climate and finding efficient ways to cope with these changes. Using various scenarios of socio-economic development and their implications for the climate, it derives optimal climate policies and determines the best time to implement these policies. The practices of climate–economic analysis strive to provide robust inference despite the computational constraints that they face and the lack of established methodology.

This dissertation focuses on development and implementation of new methods to improve conventional practices in climate–economic modeling and also on using data analysis to understand the mechanism of incorporating the information about changes in climate patterns into financial markets. Two out of the three chapters of this dissertation contribute to the integrated assessment modeling literature. Integrated assessment modeling comprises the challenge of fitting both the Earth's climate and the world economy within the computational constraints of a single model. It requires simplified yet realistic representations of the climate system and the economy—the representations that feature ensemble runs and simulations and could be embedded into an optimization framework.

Chapter 1 addresses the task of constructing such simplified representations of the climate system. In this chapter we propose a method to construct low-dimensional approximations of climate models of high complexity and demonstrate its application on the climate model MAGICC (all citations are given in respective chapters). Using a one-line statistical model we are able to approximate the climate response projected by MAGICC—the rise of global atmospheric temperature—to anthropogenic CO_2 emissions.

The core of the proposed emulation technique is the construction of uncorrelated emissions scenarios derived from orthogonal polynomials. We compare the use of the newly designed scenarios to that of the Representative Concentration Pathways (RCPs), which represent all scenarios of anthropogenic greenhouse gas emissions considered plausible in the literature. We perform simulations to show that uncorrelated scenarios are much

better suited to the task of emulation than the RCPs universally used in the literature, including approximation studies. The scenarios for these simulations are constructed in such a way that they evolve similarly to RCPs but span an even broader range of CO_2 emissions. For each such scenario the proposed model approximates the increase in atmospheric temperature very closely, with an average error of only 0.02°C, whereas training the emulator on the RCP scenarios doubles the error.

Uncorrelated scenarios feature high efficiency, in that they reduce the required number of runs of the emulated complex model. A successful emulation of MAGICC, for example, required only three model runs. The technique thereby economizes substantially on the computational resources spent on emulation.

Together with the method of emulation, we supply ready-to-use specifications of the model that can predict the atmospheric temperature anomaly driven by CO_2 emissions or, alternatively, by changes in CO_2 concentrations. This model can be directly implemented in integrated assessment frameworks as a simple yet plausible representation of the climate system, and thereby relieve the burden of high dimensionality.

We additionally examine the sensitivity of the model to the characteristics of the underlying climate processes; namely, to the assumptions about the physical processes in the climate system implied by the emulated complex model. Even when calibrated to the same initial conditions and run with the same forcing scenarios, the predictions of the complex climate models are known to diverge greatly, and these differences can undermine the accuracy of the emulating models. Our goal in this regard is to ensure that the proposed approach is not restricted to a single climate model but may potentially be applied to any of the existing ones. To do this, we test whether the method works equally well for 200 different sets of parameters obtained by combining each of the 20 climate settings with each of the 10 carbon cycle settings offered by the emulated model MAGICC. Despite some notable differences among the individual climate and carbon cycle settings, the prediction error stays very low on average and does not exceed 0.07°C even in the most extreme cases.

Chapter 2 departs from working with climate–economic models and explores observational and forecast climate data from an econometric viewpoint. As is acknowledged in the literature, even the predictions of state-of-the-art climate models exhibit systematic biases over time and space. From the economic perspective, these biases transfer into potential losses, in particular in agriculture. If the markets for the commodities sensitive to weather are efficient, the prices on these markets would capture the best knowledge available to the participants about the weather conditions for the following growing seasons. Following this intuition, in Chapter 2 we explore whether the market participants are informed

about the future changes in climate patterns over the areas where crucial crops grow.

On an example of corn and its largest producer—the US—we investigate whether the expectations on the futures market for this crop are superior (predict the weather conditions for the following growing seasons better) to those based on the best publicly available forecasts of climate models. We use climate predictions from the North American Multi-Model Ensemble database, which provides high-resolution forecasts from the major North American climate models. The spatial data is then processed so that they are suitable for relating to economic and financial data in the agricultural sector.

We use regression analysis to investigate to which extent the information incorporated into financial returns can add to the forecasts of future climate conditions issued by the ensemble of climate models and how far in advance this relationship can be detected. The results of the analysis do not allow us to claim a robust statistically significant relationship. Even though the literature suggests that the corn futures markets are efficient, the mild biases of the climate forecasts revealed by our data set cannot be unambiguously linked to financial returns.

Chapter 3 presents a general framework for robust inference from climate–economic models. These models are the frontline tools used to provide policy makers with practical guidance on timely climate actions. Yet the inference from these models has proven highly vulnerable to their initial assumptions and to the calibrated values of their input parameters. We therefore see the analysis of robustness of model inference as a critical matter that has not gained enough attention in the literature and lacks common good practice.

In this chapter we propose bringing the methods of global sensitivity analysis, which has gained wide recognition and enjoys widespread use in risk analysis, to climate economics. To demonstrate how integrated assessment modeling can benefit from such scrutiny, we apply a highly efficient method of global sensitivity analysis to the most popular model in the literature, DICE. The key feature of the method—the use of polynomial chaos expansions—enables a clear, comprehensive decomposition of the uncertainty in a model's output at very low computational costs, and hence makes the method potentially applicable to integrated assessment models of higher complexity.

Along with the methodological demonstration, the application to DICE suggests two important insights. First, sensitivity analysis of a selected subset of parameters of interest (often performed in the literature) is misleading in that it might overlook the most influential factors. Second, the opposite strategy—of pooling all existing parameters together—carries the danger of triggering spurious significance due to their internal mutual dependence. One key message of the study, therefore, is the importance of careful

analysis that fully covers the uncertainty in the model and at the same time respects its fundamental features.

Overall, the work provides both methodological and practical contributions to efficient, precise practices in the field of the economics of climate change. Chapter 1 and Chapter 3 establish new practices of more robust inference from climate–economic models—practices that economize on computational resources and contribute to a more sound basis for decision-making in the compelling challenge of changes to the climate. Chapter 2 suggests a way to utilize the information available on the financial markets along with the projections of state-of-the-art climate models to assess of the present awareness of the ongoing changes in climate patterns.

Part II

Three Essays in Economics of Climate Change

Essay 1

Statistical Approximation of Climate Models

Statistical Approximation of Climate Models[1]

Alena Miftakhova, University of Zurich

Kenneth L. Judd, Hoover Institution, Stanford University

Thomas S. Lontzek, School of Business and Economics, RWTH Aachen

Karl Schmedders, University of Zurich

Abstract

In many studies involving complex representation of the Earth's climate, the number of runs for the particular model is highly restricted and the designed set of input scenarios has to be reduced correspondingly. Furthermore, many integrated assessment models, in particular those focusing on intrinsic uncertainty in social decision-making, suffer from poor representations of the climate system due to computational constraints. In this paper we propose a general emulation method for constructing low-dimensional approximations of complex dynamic climate models. Our method uses artificially designed uncorrelated CO_2 emissions scenarios, which are much better suited for the construction of an emulator than are conventional emissions scenarios. We apply our method to the climate model MAGICC to approximate the impact of emissions on global temperature. Comparing the temperature forecasts of MAGICC and our emulator, we show that the average relative out-of-sample forecast errors in the low-dimensional emulation models are below 2 percent. Our emulator offers an avenue to merge modern macroeconomic models with complex dynamic climate models.

Note: A version of this paper has been accepted to publication in the *Journal of Econometrics*.

[1]We thank the participants at the 2014 Initiative on Computational Economics at the Hoover Institution, at the conference "Econometric Models of Climate Change 2016" at the University of Aarhus, and at the 23rd Annual Conference of the European Association of Environmental and Resource Economists (2017) for helpful comments. We are indebted to Marc Paolella for his helpful advice and to Dave Brooks for excellent editorial support. Alena Miftakhova gratefully acknowledges financial support from the Swiss National Science Foundation.

1.1 Introduction

In many studies involving complex representation of the Earth's climate, the number of runs for the particular model is highly restricted and the designed set of input scenarios has to be reduced correspondingly. Integrated assessment models (IAMs) combine dynamic models of the climate system with dynamic economic models to study their interactions and formulate policies related to limiting greenhouse gas emissions. Many IAMs, in particular those focusing on intrinsic uncertainty in social decision-making, suffer from poor representations of the climate system due to computational constraints. In this paper we address this problem by constructing a low-dimensional dynamic system that accurately represents the impact of world CO_2 emissions on world average global temperature. This is a valid reduction, because CO_2 disperses rapidly in the atmosphere and the major impact of climate change can be represented by average world temperature. The basic procedure for constructing a reduced model (often called an emulator) is to specify a set of emissions paths, use each one as input into a complex climate model, observe the resulting temperature paths, and use these simulated emissions and temperature data[2] to specify an approximating dynamic system. The structure has a pooled cross-sectional nature since each path is a time series and multiple paths are used. Given the high computational costs of running complex climate models, one important criterion for the input emissions paths is their efficiency in "extracting" information from the system. We construct uncorrelated CO_2 emissions scenarios, and show that, when used as input for simulations, they prove more efficient than conventional scenarios. In the demonstrative case of emulating a climate model, forecast errors decrease by almost half when we use uncorrelated scenarios.

IAMs, deterministic or stochastic, should use as many state variables as required to ensure a realistic specification of the climate. One commonly used climate model is the Model for the Assessment of Greenhouse-gas Induced Climate Change (MAGICC), a reduced-complexity climate emulator (Meinshausen et al., 2011a). Since the computational complexity of solving dynamic models increases with their dimensionality, a dynamic system of the size of MAGICC is too large to be commonly applied in stochastic economic models with continuous state and control variables and large time horizons. Using our approximation approach, we construct an emulator of MAGICC that produces reliable predictions of temperature response to CO_2 emissions at a much lower cost.

Although statistical approximations seem alien to the physical nature of climate models,

[2] The word "data" in econometrics conventionally means observations of the real world. In the climate change literature, in turn, "data" can refer both to historical observations and to the simulated scenarios of the future climate, as in the IPCC Data Distribution Centre (IPCC-TGICA, 2007). Following this terminology, all references to data in the present paper refer to the simulated output of the climate model concerned.

they are in fact closely connected to it. Studies show that statistical methods can produce models that adhere to basic climate physics. Kaufmann et al. (2013) test two time series models that link radiative forcing to global surface temperature as simulated by complex climate models and conclude that the model with a stochastic trend is supported more strongly by the data. They also show that this statistical model is consistent with the relation characterized by a globally averaged energy balance model, which is based on the physical relationship between radiative forcing and surface temperature. Pretis (2015) extends this analysis to a two-component (ocean–atmosphere) energy balance model by showing its equivalence to a cointegrated time series relation of the modeled climate variables.

Often, the methodology for emulating high-dimensional models assumes the preexistence of some input–output data sets from computer simulations. The general requirement for the input scenarios is that they are "carefully chosen" and "excite" the emulated system sufficiently (Young and Ratto, 2011; Castelletti et al., 2012). The existing data sets of climate models' simulations contain the predictions of complex climate models in response to some commonly prescribed scenarios (e.g., IPCC-DDC, 1998). A widely known example of such common scenarios is the Representative Concentration Pathways (RCPs) adopted in the most recent assessment report of the Intergovernmental Panel on Climate Change (IPCC, van Vuuren et al., 2011). These scenarios are designed to provide a consistent base for scientific research, from single studies to large intercomparison projects. In this paper, however, we show that they are not a wise choice for the particular task of approximation.

In cases when additional simulations are not computationally expensive, researchers generate a large collection of emissions paths and the resulting temperature paths. For example, to emulate a climate model of intermediate complexity, Holden and Edwards (2010) construct an ensemble of possible future concentration profiles using the Latin hypercube method, and apply dimensionality reduction techniques to construct an emulator. A few studies consider statistical emulation an applicable data compression tool when storing the full data sets from climate models' runs proves too costly (Castruccio and Genton, 2016; Guinness and Hammerling, 2017). With the increasing volumes of generated climate data, the methodology enables the compression of data sets while preserving their distributional properties. For such experiments it is especially important to design the runs of the emulated complex models in such a way that the resulting input–output data sets are as informative as possible.

Unfortunately, the most detailed and complex climate models are costly to run: it can take several months to simulate a few hundred model years (Dringnei et al., 2008). This limits the collection of existing simulations. If the existing data is deemed insufficient to design a robust emulator, it is expensive to significantly increase the available data set.

In a recent study, Castruccio et al. (2014) recognize the need to run more scenarios of the Community Climate System Model, version 3 (CCSM3) (Collins et al., 2006; Yeager et al., 2006). They design five new scenarios specifically for emulation and demonstrate that the trajectories of temperature and precipitation can be emulated using CO_2 concentrations paths. In choosing new scenarios, they do not follow any experimental design procedure, but make part of the scenarios similar to those readily available in multimodel experiments and let the other part induce rapid changes in CO_2 concentrations.

The key distinction of the present paper is our focus on the procedure for designing efficient input scenarios. We do not aim for our scenarios to resemble the existing runs of the complex climate models, but propose a task-driven approach to scenario creation. The computational cost of running complex climate models makes it imperative that input scenarios are chosen to maximize the information gained from these computations. This paper takes a mathematical view motivated by approximation theory. Intuitively, the input scenarios should be orthogonal in some sense. But when we apply principal component analysis (PCA) to the four conventional RCPs, we find that they jointly contain little more information than one scenario would. We, therefore, propose, instead, to use a family of orthogonal polynomials as a base for constructing the input scenarios for emulation. These scenarios obviously do not look like anything we expect will happen to the climate system; but this is not important for our task. And indeed, as intuition from approximation theory would indicate, we find that our four orthogonal polynomial input scenarios produce a significantly better emulator than the four standard RCPs. We test the usefulness of our emulation approach for stochastic IAMs by evaluating its ability to accurately simulate the distribution of temperatures in response to a stochastic emissions process. We find that it is very good at this task.

Our emulator can be directly used to improve IAMs. Currently, most IAMs are deterministic, assuming that the future climate and economy are perfectly predicable. These perfect foresight models do not entail random variability within the model but attempt to incorporate uncertainty by applying Monte Carlo simulation techniques (see, for example, Mastrandrea and Schneider, 2004 or Hope, 2013). These deterministic models are solved many times, each time with their parameters being drawn from a distribution. The approximation of the stochastic framework is then presented by probabilistically averaging the results of all deterministic simulations. These models certainly produce valuable insights, for example regarding the range of possible model outcomes. However, they do not model economic and climate-related uncertainty in the decision-maker's problem, and therefore miss real features of decision-making under uncertainty. Crost and Traeger (2013) show that the Monte Carlo approach can produce misleading implications for policy making.

Even studies that account for some kind of uncertainty often point to the "curse of dimen-

sionality" as an explanation for their simplified representations of the climate (EPA, 2010; Webster et al., 2012; Newbold et al., 2013; Jensen and Traeger, 2014). Moreover, stochastic IAMs often rely on very simple assumptions when formulating the actual stochastic processes.[3] However, recent advances in computational methods have described ways to solve high-dimensional economic models, even beyond 100 dimensions (Maliar and Maliar, 2015; Judd et al., 2011; Cai et al., 2015; Brumm and Scheidegger, 2017). Combining our approach to the construction of emulators with these new computational methods offers the potential to build more realistic models.

The remainder of the paper is organized as follows: in Section 1.2 we describe the methods used for approximation; we motivate the use of uncorrelated scenarios, explain the procedure of constructing them, and present the general model we estimate. Based on the results of the estimations, Section 1.3 states the specifications for recommended representations of the climate system and assess our emulator's performance on alternative scenarios and under different sets of initial parameters of the climate model. Section 1.4 concludes.

1.2 Approximation of high-dimensional models

In this section we describe our approach to the construction of an emulator of a complex climate model. First, we briefly describe our source of temperature predictions, MAGICC. Second, we explain why conventional emissions scenarios are not a good source for the construction of an emulator. Next, we derive a set of uncorrelated emissions scenarios, which is ideally suited for the derivation of an emulator. And finally, we depict our general constructive approach.

1.2.1 MAGICC model

MAGICC is a carbon cycle–climate model used to emulate the insights from large and complex Atmosphere–Ocean General Circulation Models (AOGCMs). The emulation task is to generate global temperature responses (the output variable) to various exogenous emissions scenarios (the input variable). In the climate research community, MAGICC is considered to be a reduced complexity climate model. Yet, it includes representations of the most essential physical and biological components of complex AOGCMs.

The emulation is performed in four steps shown in Figure 1.1. Each of the four modules contains its own set of equations; the total number of involved parameters exceeds 400. The first module of MAGICC includes terrestrial and ocean carbon cycles, which infer the

[3]Cai et al. (2015) and Lontzek et al. (2015), among others, discuss the implications of structural assumptions in IAMs regarding modeling the risk of abrupt and irreversible catastrophic climate events.

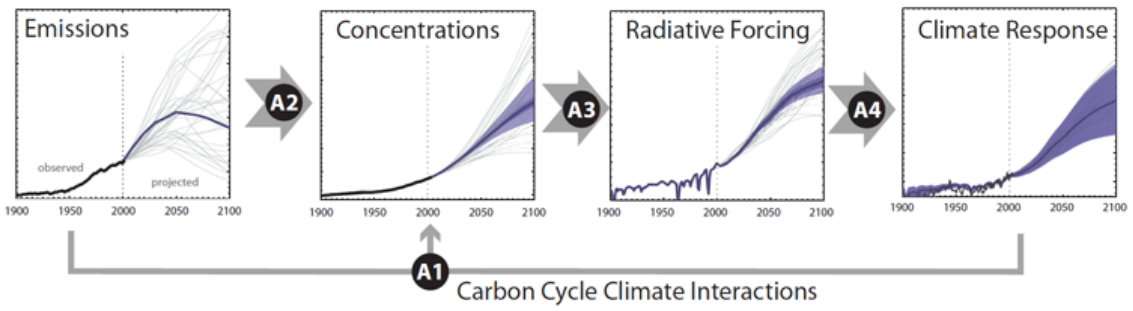


Figure 1.1: Schematic overview of the MAGICC model showing the key steps of inferring climate response from emissions scenarios. The labels $A1$ to $A4$ refer to the corresponding sections in the Appendix A in Meinshausen et al. (2011a).
Source: Fig A.1. in Meinshausen et al. (2011a); retrieved from `http://wiki.magicc.org`.

concentrations of CO_2 and other forcing gases in the atmosphere based on their individual cycles and atmospheric life-times. For example, the change in CO_2 concentrations is induced by anthropogenic emissions and ocean and terrestrial fluxes (net uptakes or releases of CO_2 by ocean and land) calculated in separate box models. The changes in concentrations are subsequently transformed into induced radiative forcings (net flows of energy into the atmosphere) of the gases according to assumed approximating relationships.[4] Finally, the changes in radiative forcing are aggregated over the two hemispheres and transfer into the change in global average temperature level via a global energy-balance equation.[5]

Despite being a "simple" model, MAGICC performs exceptionally well in emulating the results of the large AOGCMs (see Meinshausen et al., 2011a). MAGICC was therefore used in the recent IPCC reports (Pachauri et al., 2014) as the prime tool for evaluating carbon and climate responses to various emissions scenarios. MAGICC is publicly available and easy to operate. According to Meinshausen et al. (2011a), MAGICC is flexible enough to deliver accurate results when running scenarios outside of the original calibra-

[4]For example, following the IPCC convention, CO_2 radiative forcing, $\triangle R_{CO_2}$, is assumed proportional to the logarithm of the ratio of the current CO_2 concentration, S, to its preindustrial level, S_0,

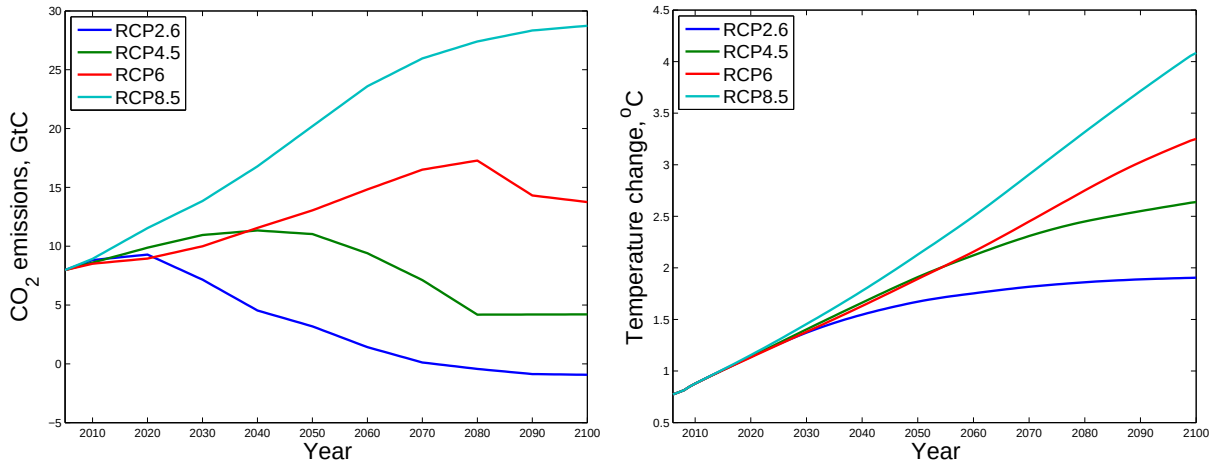
$$\triangle R_{CO_2} = \alpha \log(S/S_0),$$

with the latest estimate of the value of α equal to $5.35 W m^{-2}$ (Myhre et al., 1998). Radiative forcings of methane, nitrous oxide, and other gases are modeled via similar relations.

[5]The global mean radiative forcing comprises atmospheric and oceanic energy fluxes. The general form of the energy-balance equation,

$$\triangle R_G = \lambda_G T_G + \frac{dH}{dt},$$

relates the global mean radiative forcing, $\triangle R_G$, and the temperature anomaly, T_G, through the feedback factor, λ_G. The last term in the equation denotes the ocean energy flux and is assumed equal to zero in equilibrium state. The energy balance equations for both hemispheres additionally take into account ocean and land feedback factors as well as land-ocean and hemispheric heat exchanges.



(a) CO_2 emissions, RCP scenarios. (b) Temperature anomaly, RCP scenarios.

Figure 1.2: RCP emissions scenarios (a) and corresponding predictions of temperature (b).

tion space. For the purpose of the present paper, MAGICC is, therefore, best suited to generating reliable responses of global temperature to any emissions scenario—responses that we use as benchmarks for evaluating the accuracy of our statistical model.

We chose the model MAGICC because its computer implementation is very fast (a matter of seconds on many computers at the time of writing of this paper). This speed of implementation makes it feasible for us to compare the temperature forecasts of our emulator to those of MAGICC for 200 different alternative model settings (see Section 1.3.4). In addition, even for such a fast climate model an emulator may make it much easier to incorporate the model's insights into a complex macroeconomic model.

1.2.2 Conventional emissions scenarios

The four basic RCPs endorsed by the IPCC are the most common and ready-to-use scenarios and currently serve as a coherent base for integrated climate–economic modeling and model intercomparison projects (Taylor et al., 2012). Each RCP is a distinct pathway of radiative forcing named after the level of forcing it achieves in 2100. RCP2.6, RCP4.5, RCP6, and RCP8.5 thus imply forcing of 2.6 to 8.5 W/m^2 by the end of the century. Each scenario also specifies the associated levels of emissions and concentrations of greenhouse gases and other forcing agents. Figure 1.2 displays the decadal CO_2 emissions paths implied by the RCPs and the corresponding temperature rises as predicted by MAGICC.[6] The lowest-forcing scenario is the only one that does not hit the critical value of warming the atmosphere to 2℃ above the preindustrial level; the most aggressive scenario among the four—RCP8.5—implies warming of above 4℃.

[6]The RCP scenarios are specified on a decadal scale; MAGICC uses their linear interpolation to produce the corresponding annual paths of temperature anomalies.

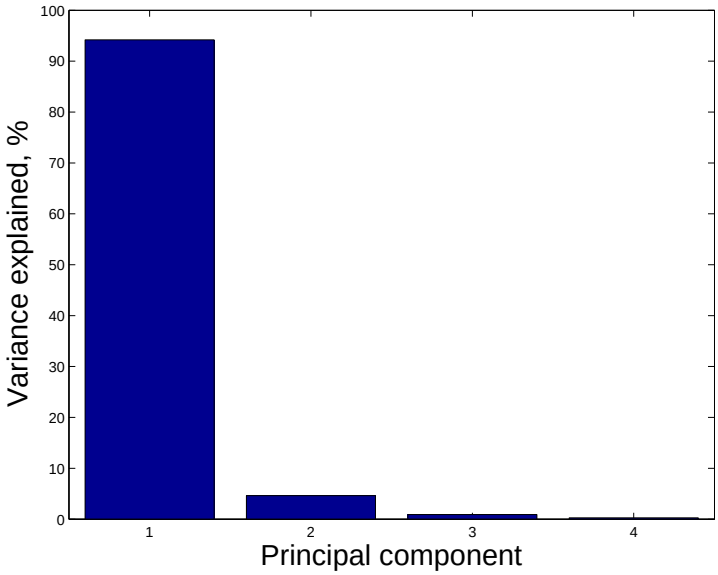


Figure 1.3: Variance decomposition from principal component analysis of the RCP scenarios.

The RCPs are an example of a scenario set designed to provide reference pathways to scientists, integrated assessment modelers, and policy makers (Moss et al., 2010). They were selected to represent scientific agreement with regard to the probable trajectories of future climate and socioeconomic conditions. It is not the ultimate goal of such universal scenarios to serve as an input for approximations. We show below that common scenarios such as the RCPs are not suitable for estimating model parameters, and that the shape of scenarios should be determined by the purpose of their use.

Principal component analysis indicates that the RCPs are not likely to be a good set when it comes to estimating an emulator. Figure 1.3 displays the variance decomposition implied by PCA, and shows that the first principal component carries more than 94 percent of the total variance in the set; the first principal component, which is a linear combination of the four RCPs, provides nearly as much information about emissions as the four scenarios do collectively.

The statistical approach to emulator construction uses the emissions–temperature pairs to estimate a single time series model where temperature is the dependent variable and lagged dependent and independent variables are on the right-hand side. The structure is similar to that of a pooled cross-section problem. Even though we use terms like "statistical approach", it needs to be emphasized that there is no underlying stochastic structure to the problem. The problem is really one of approximation, where we want to find a simple dynamic model relating temperature to emissions with small prediction errors.

1.2.3 Orthogonal emissions scenarios

To motivate our choice of emissions scenarios for the construction of an emulator, it is helpful to take a closer look at the underlying problem. Suppose some scientists or policy makers would like to consider an additional scenario of CO_2 emissions, $E(t)$, for years $t = t_0, t_1, \ldots, t_L$. Unfortunately, they cannot quickly access a complex dynamic climate model to forecast the resulting temperature anomalies, and therefore must resort to an emulator. Such an emulator relies on inputs from certain previous emissions scenarios, $E_k(t)$, $k = 0, 1, \ldots, m$, for which temperature predictions were gathered from that aforementioned complex model. But in which cases will an emulator trained on some existing scenarios, $E_k(t)$, provide a good approximation for the new scenario, $E(t)$?

We argue that the most likely or most popular emissions scenarios, such as the four afore-mentioned RCPs, are not necessarily a good set of input scenarios for training an emulator, because such scenarios are likely to be strongly correlated. (The four RCPs are a case in point.) Instead, a desirable condition of a good training set of emissions scenarios is that they span as large a set of reasonable scenarios as possible. Put differently, for any reasonable future scenario we would like there to exist weights, a_k, $k = 0, 1, \ldots, m$, such that

$$E(t) \approx \sum_{k=0}^{m} a_k E_k(t) \tag{1.1}$$

for all t.

The first of several questions that naturally arise is whether such an approximation of a new scenario by using previous scenarios is even possible. An answer in the affirmative will require that the set of scenarios employed for the approximation offers sufficient flexibility and scope to approximate any new (reasonable) scenario. RCP scenarios such as those depicted in Figure 1.2(a) are (discrete) time series of CO_2 emissions with fixed start and end dates. Therefore, a natural method of approximating them is a least squares regression approach using a suitable set of basis vectors.

Trefethen (2013) provides an excellent introduction to the approximation of one-dimensional functions in theory and practice. He strongly advocates interpolation methods relying on Chebyshev polynomials for the practical approximation of functions on intervals, and literally refers to such methods as "unbeatable". Moreover, this particular family of orthogonal polynomials appears also to be the most popular choice for approximation problems in which we want to use more points than the maximum order of the Chebyshev polynomials. When the number of conditions exceeds the number of basis elements we need to resort to a regression approach.

In this paper we use Chebyshev polynomials to create basis vectors of CO_2 emissions. These newly created scenario vectors, E_k, are then used as explanatory variables in a

regression with a new emissions scenario vector, E, as the dependent variable. If the scenario vectors E_k are a good basis for the approximation of an arbitrary scenario E, they can also be a good input set for creating an emulator that ultimately needs to provide forecasts for that new scenario.

Chebyshev polynomials and regression

For an excellent treatment of Chebyshev polynomials we refer interested readers to Trefethen (2013). Here, we only provide a brief introduction to Chebyshev polynomials, which suffices for our purposes.

Without loss of generality, consider the interval $[-1, 1]$. Consider the following recurrence relation:

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_k(x) = 2xP_{k-1}(x) - P_{k-2}(x) \text{ for } k = 2, 3, \dots .$$

The polynomial[7] P_k is called the kth *Chebyshev polynomial*. The Chebyshev polynomials P_0 through P_7 are as follows:

$$\begin{aligned} P_0(x) &= 1 & P_4(x) &= 8x^4 - 8x^2 + 1 \\ P_1(x) &= x & P_5(x) &= 16x^5 - 20x^3 + 5x \\ P_2(x) &= 2x^2 - 1 & P_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1 \\ P_3(x) &= 4x^3 - 3x & P_7(x) &= 64x^7 - 112x^5 + 56x^3 - 7x. \end{aligned}$$

For each $k \geq 1$, P_k is a polynomial of degree k with a leading coefficient 2^{k-1}; it satisfies $-1 \leq P_k(x) \leq 1$ for $x \in [-1, 1]$ and it has the k zeros

$$x_n = \cos\left(\frac{(2n+1)\pi}{2k}\right) \quad \text{for } n = 0, 1, \dots, k-1,$$

which all lie in the interval $[-1, 1]$. These zeros are also called *Chebyshev nodes*.

For our study, an important property of Chebyshev polynomials is their discrete orthogonality property. If $\{x_n\}_{n=0}^{k-1}$ are the k zeros of P_k, then for all $i, j \leq k$

$$\sum_{n=0}^{k-1} P_i(x_n)P_j(x_n) = \begin{cases} 0 & : i \neq j, \\ k & : i = j = 0, \\ k/2 & : i = j \neq 0. \end{cases}$$

So, for $i \neq j$, the vectors of the values of P_i and of P_j at the Chebyshev nodes $\{x_n\}_{n=0}^{k-1}$ are orthogonal and thus uncorrelated.

[7] We deviate from the standard notation and denote Chebyshev polynomials by P instead of by T or C. We need the letters T and C for the time series of temperatures and cumulative CO_2 emissions, respectively.

All Chebyshev zeros lie in the interval $[-1, 1]$. For an approximation on a general interval, $[a, b]$, we use the standard linear transformation,

$$t = \left(\frac{b-a}{2}\right) x + \frac{a+b}{2}, \tag{1.2}$$

for $t \in [a, b]$ and $x \in [-1, 1]$.

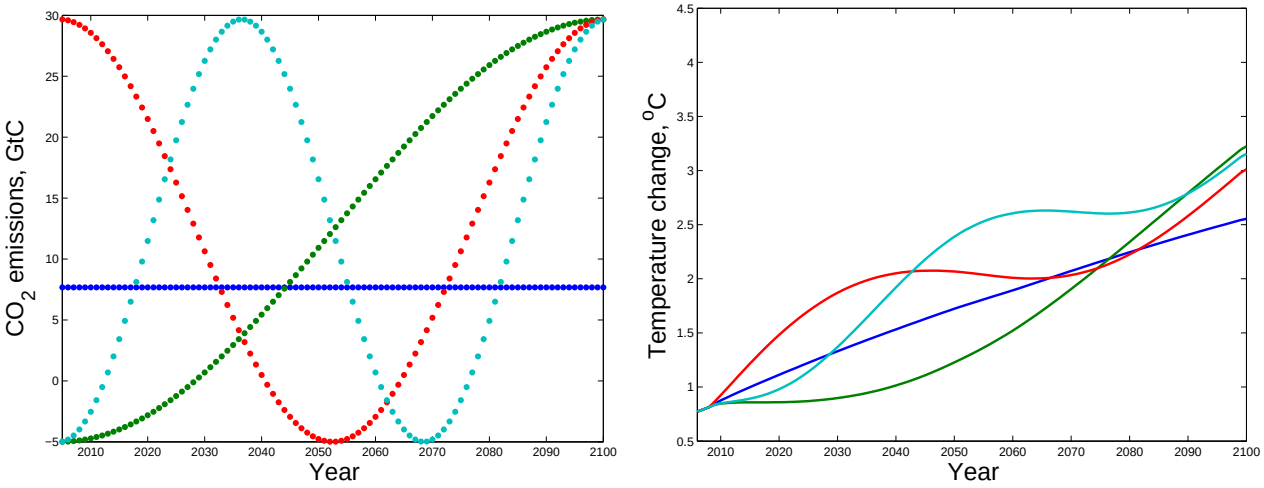
To apply the standard Chebyshev regression approach, we could now proceed as follows: First, choose a number of n Chebyshev nodes—that is, all zeros of P_n—and evaluate the Chebyshev polynomials of degrees 0 to m at these n points in order to obtain $m + 1$ basis vectors of length n. Second, determine the values of the to-be-approximated new scenario at the n nodes. Since the n nodes would not coincide with integer values—the years $2005, 2006, \ldots, 2100$—we would need to use interpolated values from the new scenario. Third, determine the regression coefficient by minimizing some criterion—for example, the sum of squared residuals.

We do not pursue this standard approach because it encounters the following problem: Both the input scenarios and the temperature output of a typical climate change model are annual figures—that is, they are given for the years $2005, 2006, \ldots, 2100$. Therefore, we must determine the values of the Chebyshev polynomials at these integer points, instead of at the Chebyshev nodes, when we want to enter the corresponding base scenario into a model such as MAGICC. But then the adjusted values of the base vectors at the integer points are no longer uncorrelated. While some small correlation among the base scenario vectors may do only little harm to the regression, we would like to avoid it if possible. Therefore, we generate the data for the regression approach in a slightly different way.

Design of uncorrelated scenarios

The first step of our regression approach is identical to that of the general approach described above. We choose n to be the number of years in our simulated data; here $n = 96$, since we consider scenarios for the time period 2005–2100. For the n zeros of P_n, we evaluate the Chebyshev polynomials of degrees 0 to m at these n points in order to obtain $m + 1$ basis vectors of length n. By construction, these $m + 1$ basis vectors are pairwise uncorrelated. Now we treat the ith element of each basis n-vector as the value for the year $2004 + i$. As a result, in the second step, we no longer need to interpolate values for the new scenario but can just take the given scenario values. Using this different vector as the independent variable we can determine regression coefficients and check whether the new scenario can be approximated well.

When we determine and evaluate the linear approximation, $\sum_{k=0}^{m} a_k E_k(t)$, for a new sce-



(a) CO_2 emissions, uncorrelated scenarios. (b) Temperature, uncorrelated scenarios.

Figure 1.4: Uncorrelated emissions scenarios (a) and corresponding predictions of temperature (b).

nario, $E(t)$, the transformation of the Chebyshev polynomials to the domain $[2005, 2100]$ suffices and we do not need to adjust the range value. However, before we can feed the corresponding emissions scenarios into a model such as MAGICC, we also need to linearly transform the range of the polynomials in order to obtain CO_2 emissions scenarios of reasonable sizes. Since the four RCPs in Figure 1.2(a) span a range of 29.6708℃, we use (1.2) to transform the m scenarios based on the Chebyshev polynomials of degrees 1 to m to the range $[-5, 29.6708]$. The zero-degree polynomial is an exception: it corresponds to a scenario of constant annual emissions, or a steady state of the economy, and is therefore set to the last historical value of CO_2 emissions.

The resulting values represent CO_2 emissions levels for the $n = 96$ years, 2005–2100. For the baseline case for our analysis, we set $m = 3$ so that using four uncorrelated scenarios can be compared to using the four RCPs. Figure 1.4(a) depicts the resulting CO_2 emissions scenarios. When we enter them into MAGICC we must also specify scenarios for some other gases. For simplicity, in this study we set the annual emissions of all other gases in the years 2005–2100 to their average levels across the four RCP scenarios. The emissions of all gases for all years prior to 2005 are kept at their historical values. Figure 1.4(b) shows the resulting temperature anomalies forecasted by MAGICC for the four uncorrelated scenarios.

Figure 1.5 shows the variance decomposition from a principal component analysis of the four base scenarios. By construction, each of the first three components explains a third of the variance. The degree-zero polynomial (a constant) obviously does not carry any variance but is nevertheless included in the set of the base scenarios. Recall that a good set of input scenarios for emulation should provide a base for the linear approximation

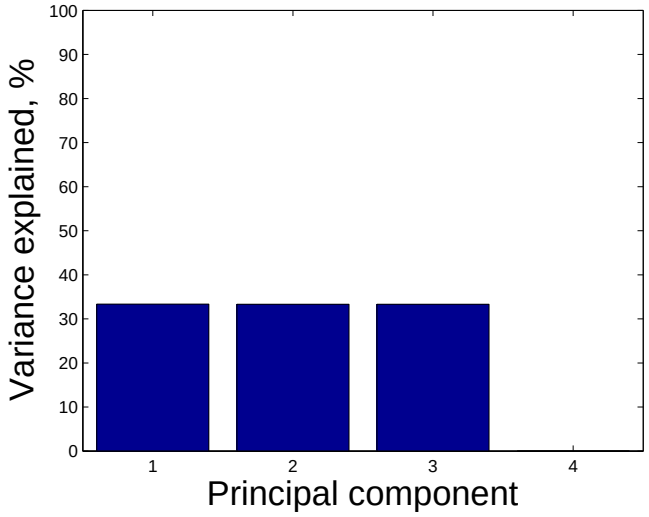
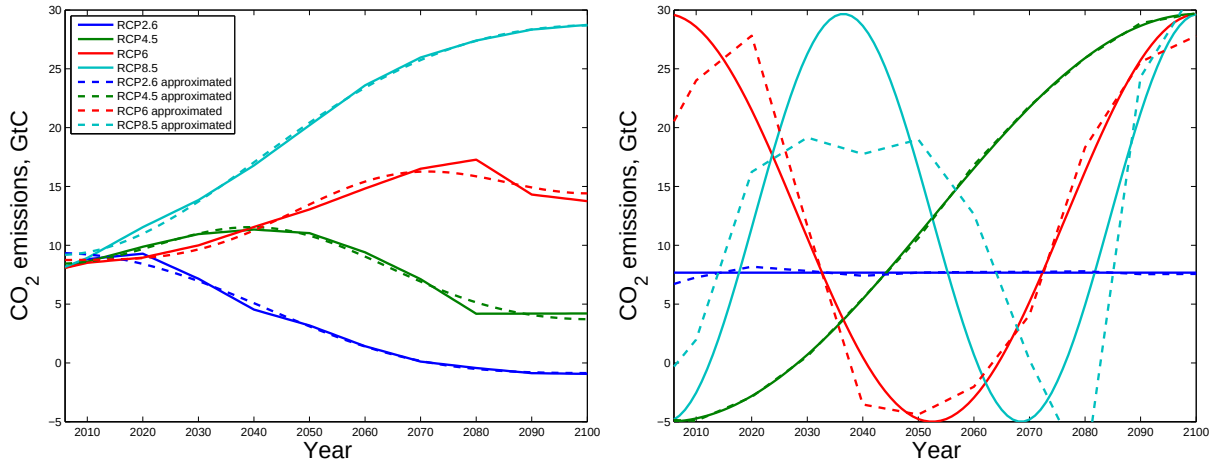


Figure 1.5: Variance decomposition from principal component analysis of the four uncorrelated scenarios.

(1.1). For a successful linear approximation, the set of basis vectors has to include a constant.

The set of base scenarios, besides carrying the zero correlation property, should trigger a strong enough response in the approximated system. The advantage of the scenarios based on the Chebyshev polynomials is that they include the chosen extreme values of emissions. Each Chebyshev polynomial of degree $k \geq 1$ has $k + 1$ extrema, with their values at the endpoints of the range $[-1, 1]$. Each corresponding emissions scenario scaled to a range $[a, b]$ reaches the minimum and maximum emissions levels a and b at those extreme points. This property ensures that we observe the simulated response of the climate system when the CO_2 emissions levels reach their extreme values, at a slower or faster rate.

Clearly, future emissions paths will not look anything like the base scenarios based on the degree two or degree three Chebyshev polynomials. But the task here is not to discuss what the likely emissions paths are. The task is to find a collection of emissions scenarios that enable us to extract as much information as possible from a mathematical model of the climate such as MAGICC. The key property of the uncorrelated base scenarios is that they produce good approximations of the standard RCP scenarios, which are of particular interest in the literature. Fig. 1.6(a) demonstrates that the designed uncorrelated scenarios, while unrealistic, provide an excellent basis for close approximations of the RCPs; in fact, the root-mean-squared error (RMSE) of the approximation is only 0.36 gigatons of carbon per year ($GtCyr^{-1}$). The RCPs, designed to represent likely paths, do not form a suitable collection for emulator construction. Figure 1.6(b) shows that they do not approximate the uncorrelated base scenarios well, and produce an RMSE of



(a) Approximation of RCP scenarios. (b) Approximation of base scenarios.

Figure 1.6: Approximation of the RCP scenarios with uncorrelated base scenarios (a) and approximation of base scenarios with the RCP scenarios (b).

$2.55 GtCyr^{-1}$. This result is consistent with PCA, which shows that the four standard RCPs together contain little more information than the best one on its own.

1.2.4 Construction of an emulator

With the uncorrelated emissions scenarios at hand, we can proceed to building an emulator of a climate model of our choice. To generate the data for emulation, we run the four designed emissions scenarios in MAGICC and collect its temperature forecasts for each of them. The particular structure of an emulator is defined by its purpose and potential application. In this paper we provide an example of an emulator that can be first of all used in integrated assessment modeling, where dynamic models need a simple way of producing forecasts for temperature rise if they are to assess its feedback effect on the economy.

Our emulator takes the following general form of a dynamic linear model:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \sum_{j=2}^{J} \beta_j X_{j,t-1} + \varepsilon_t, \tag{1.3}$$

where Y_t is the predicted variable in year t, $X_{j,t}$ is the jth covariate in year t, and J is the number of covariates in the model. To capture the residual autocorrelation, we assume that model errors follow an ARMA(1,1) process and that its residuals are normally distributed,

$$\varepsilon_t = a\varepsilon_{t-1} + bu_{t-1} + u_t, \; u_t \overset{iid}{\sim} N(0, \sigma^2). \tag{1.4}$$

The temperature anomaly forecasts produced by MAGICC are the benchmark values of

the predicted variable Y. The candidates for covariates come from the input scenarios and depend on the ultimate goal of the emulation. An emulator that predicts temperature response to CO_2 emissions can serve as a substitute for the entire climate module in an integrated assessment model. Alternatively, an IAM can preserve its carbon cycle representation as a distinct feature—in this case, emulation covers the transition from concentrations to temperature rise. Consequently, we consider CO_2 emissions, cumulative CO_2 emissions, and CO_2 concentrations as the available set of variables to include in X.

For each potential set of covariates, we pool the simulated data from the four runs of the climate model together and find the best values for the parameters using a maximum likelihood estimator. To test our model, we take the CO_2 emissions scenarios given by the four RCPs—they constitute our out-of-sample testing set. We use the average RMSE of predictions for this out-of-sample set of scenarios to assess the quality of their predictions.

1.3 Results

Using different combinations of the available covariates, we choose an exact form for the emulator (1.3). This section presents the results for the best two- and three-dimensional models, according to the quality of their out-of-sample predictions.

1.3.1 Best-performing low-dimensional model

In all our specifications, the dependent variable is the temperature anomaly T with respect to the preindustrial temperature level—that is, the increase in temperature since 1765. We find that the following model specification with cumulative CO_2 emissions, C_t, as a single exogenous variable produces the best predictions:

$$T_t = \beta_0 + \beta_1 T_{t-1} + \beta_2 C_{t-1} + \varepsilon_t, \tag{1.5a}$$

$$\varepsilon_t = a\varepsilon_{t-1} + bu_{t-1} + u_t,\ u_t \overset{iid}{\sim} N(0, \sigma^2). \tag{1.5b}$$

Here, cumulative CO_2 emissions, C_t, are accumulated from the year 1765 to t and measured in $GtC \cdot 10^3$. The scenarios themselves provide values for the lagged cumulative emissions starting in 2005. For each uncorrelated emissions scenario, we obtain temperature anomalies from MAGICC for $t \in \{2005, 2006, \ldots, 2100\}$. And so, the number of data points for the estimation of this model is 96 times the number of scenarios, m. The first row of Table 1.1 reports the estimated values for the parameters of model (1.5) and the model's average errors of out-of-sample prediction. The average RMSE across all four testing scenarios is only about 0.03℃; the prediction errors are on average 1.62% of the level of temperature rise.

Table 1.1: Approximation results for different model specifications. The last two columns report the average errors of prediction for the testing set of the RCP scenarios. RMSE is the root-mean-squared error; MAPE is the mean absolute percentage error. The estimated standard errors of the coefficients are given in parentheses.

Model	β_0	β_1	β_2	β_3	a	b	σ	RMSE	MAPE
(1.5)	0.2500	0.7650	0.3632		0.9805	0.2128	0.0022	0.0338	1.62%
	(0.0344)	(0.0227)	(0.0345)		(0.0031)	(0.0636)	(0.0001)		
(1.6)	0.1188	0.6874	0.1503		0.9878	0.1854	0.0018	0.0426	1.75%
	(0.0401)	(0.0193)	(0.0089)		(0.0012)	(0.0558)	(0.0001)		
(1.7)	0.1230	0.6820	0.0286	0.1445	0.9873	0.1832	0.0018	0.0411	1.68%
	(0.0396)	(0.0209)	(0.0399)	(0.0121)	(0.0013)	(0.0572)	(0.0001)		

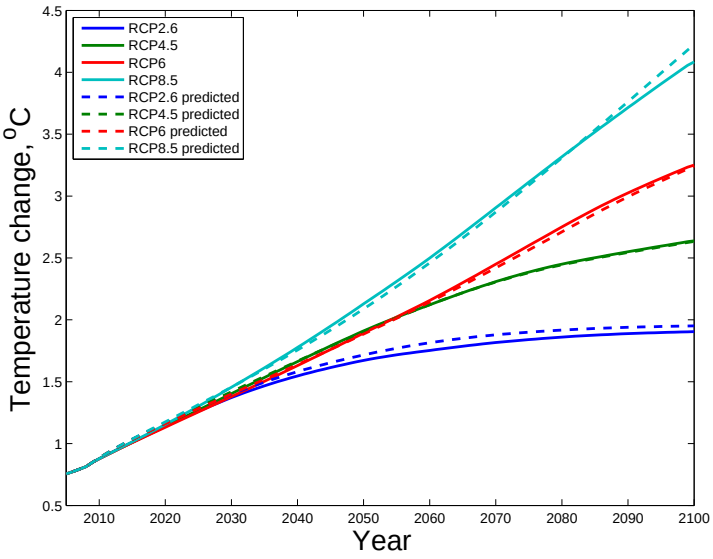


Figure 1.7: Out-of-sample temperature predictions of the best-fitting model (1.5).

Figure 1.7 demonstrates that the model produces very accurate predictions. With cumulative emissions being the best candidate for an exogenous variable, our results are consistent with recent studies that suggest a linear-proportional relationship between global warming and the level of cumulative CO_2 emissions (Allen et al., 2009; Zickfeld et al., 2009; Matthews et al., 2009; Alexander et al., 2013). In other words, even when only CO_2 emissions data are available, the temperature anomaly can be approximated instantly, leaving out the emulation of the carbon cycle and any other intermediary steps.

The required number of scenarios becomes a very important question when complex, computationally expensive models, such as AOGCMs, are emulated. Our approach suggests that emulation does not necessarily benefit from running the complex model as many times as possible. Figure 1.8 demonstrates that only up to three runs contribute significantly to the precision of the approximation, after which the prediction errors level off. The key to successful emulation in such a case is that the input data set is efficiently designed before the emulated climate models are run. Furthermore, our recommended

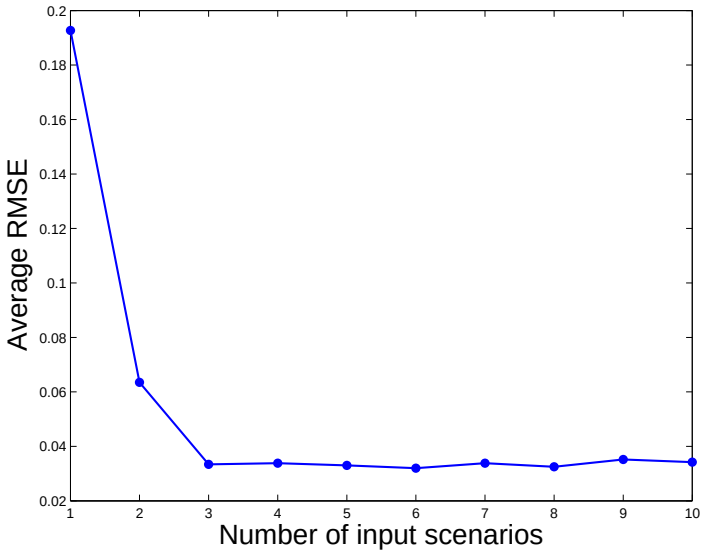


Figure 1.8: Average out-of-sample prediction error as a function of the number of uncorrelated scenarios used as input.

prediction model can be embedded within a dynamic system of equations, such as is often employed in integrated assessment models. Many IAMs—in particular those focusing on intrinsic uncertainty in social decision-making—suffer from poor representations of the climate system due to computational constraints. Given the accuracy of our predictions, integrated assessment modelers may now include this simple yet accurate low-dimensional mapping of emissions to temperature levels in their models.

Given the structure of the two dynamic equations from above, $T_t = f(T_{t-1}, C_{t-1}, \varepsilon_t)$ and $\varepsilon_t = g(\varepsilon_{t-1}, u_{t-1}, u_t)$, an economic IAM would also require adding a dynamic equation for cumulative emissions, $C_t = h(C_{t-1}, E_t)$, where E_t is some emissions scenario resulting from the model's economic framework.

1.3.2 Alternative specifications

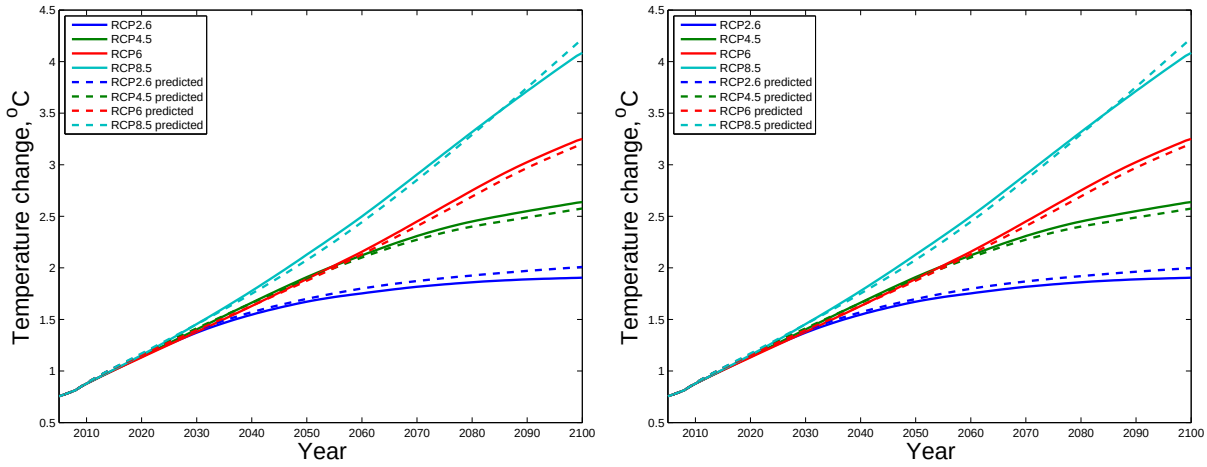
Here, we present alternative functional forms of two-dimensional and three-dimensional representations of the climate system.

The alternative two-dimensional model includes CO_2 concentrations, S_t, measured in $ppm \cdot 10^2$ in year t as an exogenous covariate,

$$T_t = \beta_0 + \beta_1 T_{t-1} + \beta_2 S_{t-1} + \varepsilon_t, \tag{1.6a}$$

$$\varepsilon_t = a\varepsilon_{t-1} + bu_{t-1} + u_t, \; u_t \overset{iid}{\sim} N(0, \sigma^2). \tag{1.6b}$$

The results reported in Table 1.1 show that cumulative CO_2 emissions are a better predictor; however, if only CO_2 concentrations are available, the resulting emulator would



(a) Two-dimensional model (1.6). (b) Three-dimensional model (1.7).

Figure 1.9: Out-of-sample temperature predictions of the alternative models.

also generate good predictions.

The suggested three-dimensional model includes both CO_2 concentrations and cumulative emissions,

$$T_t = \beta_0 + \beta_1 T_{t-1} + \beta_2 S_{t-1} + \beta_3 C_{t-1} + \varepsilon_t, \tag{1.7a}$$

$$\varepsilon_t = a\varepsilon_{t-1} + bu_{t-1} + u_t, \; u_t \overset{iid}{\sim} N(0, \sigma^2). \tag{1.7b}$$

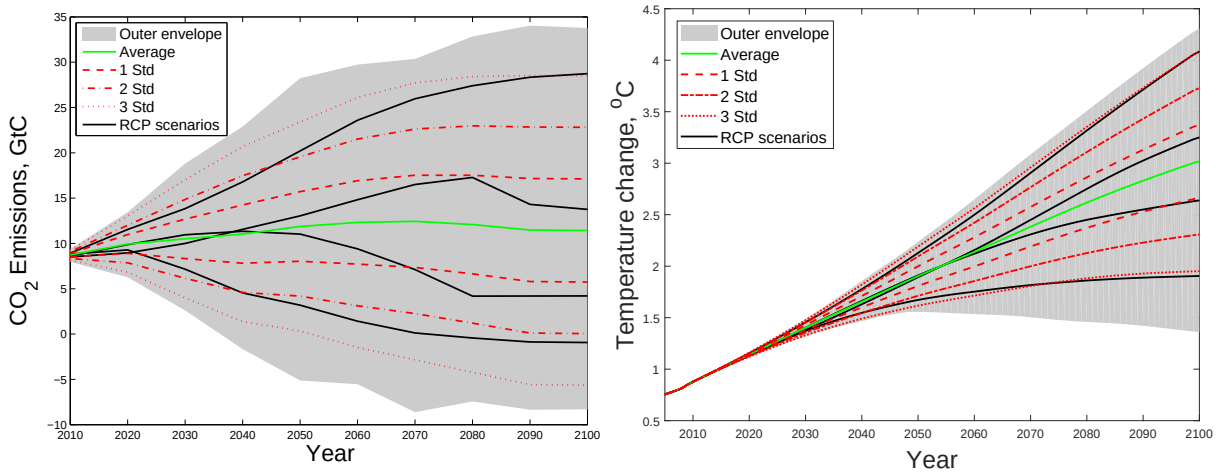
The results in the last row of Table 1.1 indicate that the extended model does not outperform the two lower-dimensional ones.

Figure 1.9 shows the out-of-sample predictions of temperature levels produced by the two alternative models. Visually, the approximations delivered by the alternative models look similar to those of the best model (see (1.5)).

1.3.3 Performance verification

Advances in climate and economic research bring with them new knowledge about likely paths of socioeconomic development and estimated climate impacts. An ongoing process of scenario creation addresses the research community's need for new scenarios that are consistent with the current understanding of possible global developments and that reflect the associated uncertainty (Moss et al., 2010). The most recent example of such is the concept of Shared Socioeconomic Pathways, which complement existing scenarios by including various paths of socioeconomic development (Riahi et al., 2017; O'Neill et al., 2014).

As new generations of scenarios are incorporated into integrated assessment modeling, they become a new common base for scientific research in this field. However, models



(a) CO_2 emissions scenarios. (b) Temperature anomaly.

Figure 1.10: Statistical distributions of the simulated scenarios (a) and of the corresponding temperature predictions of MAGICC (b).

trained on traditionally used scenarios, such as the RCPs, might perform poorly as these new scenarios come into play.

It is therefore important for us to ensure that the proposed emulator works on the entire range of scenarios considered plausible in the literature. Since the RCPs were created to represent the wide range of scenarios present in the prior literature, we would like to assess the performance of our model on this range of scenarios. In particular, we construct a stochastic process that allows us to generate a large testing set of CO_2 emissions paths similar in their nature to those present in the RCPs. We then verify whether the temperature levels predicted by our emulator for the simulated scenarios are close to those produced by MAGICC.

In principle, the following simple stochastic process can generate the decadal CO_2 emissions paths of the four RCP scenarios:[8]

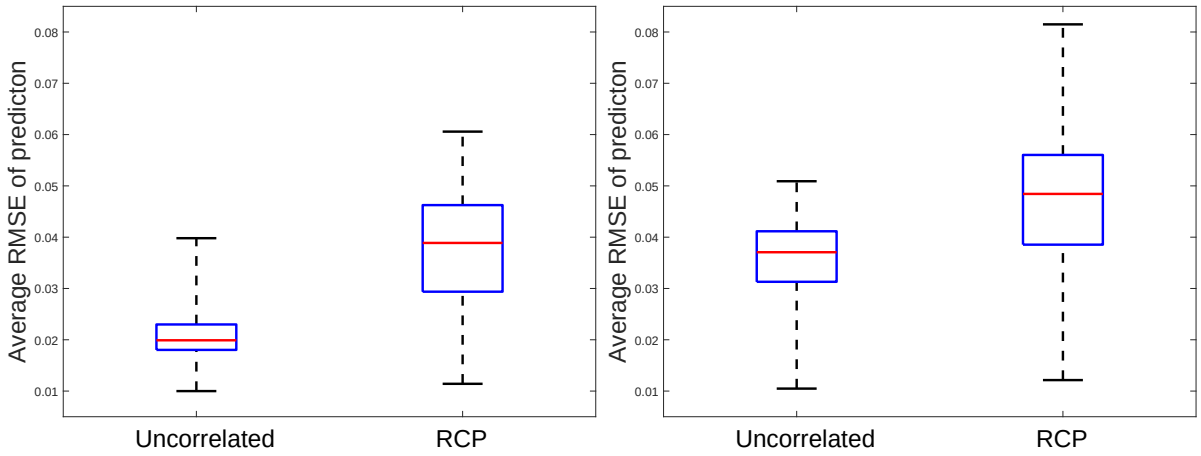
$$E_\tau = E_{\tau-1} + \varepsilon_\tau, \ \varepsilon_\tau \sim N(\mu_\tau, \sigma_\tau), \tag{1.8}$$

where E_τ is the annual CO_2 emissions level in decade τ. The parameters μ_τ and σ_τ can be estimated from the four RCP scenarios for each τ. Using the method of moments,

$$\hat{\mu}_\tau = \frac{1}{n}\sum_{i=1}^{n} \hat{\varepsilon}_{i,\tau}, \quad \hat{\sigma}_\tau = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\hat{\varepsilon}_{i,\tau} - \hat{\mu}_\tau)^2}, \text{ for } \tau = 2010, 2020, \ldots, 2100,$$

where $\hat{\varepsilon}_{i,\tau} = E_{i,\tau} - E_{i,\tau-1}$ and $n = 4$ for each decade τ.

[8]Because the RCP scenarios are specified on a decadal scale, we first generate decadal scenarios and then linearly interpolate them to use the resulting annual paths.



(a) Model (1.5) with cumulative emissions. (b) Model (1.6) with concentrations.

Figure 1.11: Out-of-sample prediction errors of models (1.5) (a) and (1.6) (b) for simulated scenarios in the cases of using covariates from uncorrelated scenarios and RCP scenarios for training the models.

We generate a testing set of 10,000 realizations of the stochastic process (1.8) (Figure 1.10) and run these realizations in MAGICC to obtain the benchmark predictions of temperature levels. We then compare the predictions of our emulator with the benchmark values in the cases of training the emulator on the uncorrelated scenarios and on the RCP scenarios. In the case of using the uncorrelated scenarios as a training set and the best-fitting model (1.5) as an emulator, we obtain an average RMSE of only 0.02℃. The corresponding error in the case of using the RCP scenarios is almost twice as big. A statistical comparison of the performance of the uncorrelated scenarios with that of the RCP scenarios as input sets for models (1.5) and (1.6) (Figure 1.11) confirms that the average prediction error decreases significantly when the designed input paths are uncorrelated.

1.3.4 Performance on alternative model settings

The predictions of complex climate models are known to diverge greatly from one another—even when calibrated to the same observational data and run with the same forcing scenarios they span a wide range of possible system forecasts. These discrepancies might stem from differences in assumptions, modeled components, and the structure of those components (Tebaldi and Knutti, 2007a), and can have a significant impact on the accuracy of emulating models (Meinshausen et al., 2011a).

So far, we have used in our analysis only one (default) combination of the 20 AOGCMs and 10 carbon cycle models emulated in MAGICC (Meinshausen et al., 2011b). However, as documented by numerous intercomparison projects, there is great variability among complex models in terms of their predictions of climate response to emissions scenarios

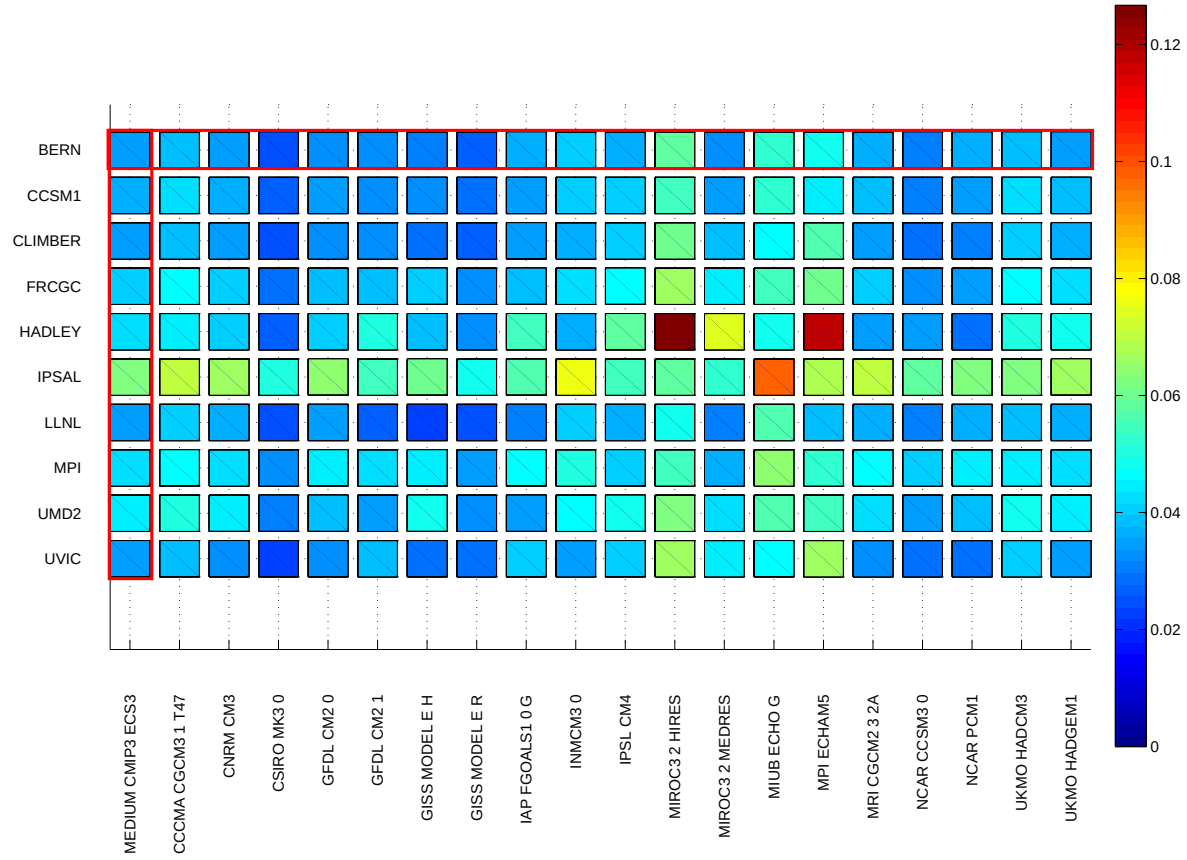


Figure 1.12: RMSEs for combinations of 20 AOGCM settings (horizontal axis) and 10 carbon cycle settings (vertical axis). The settings used in MAGICC as default values are marked with red rectangles.

(Taylor et al., 2012; Palmer et al., 2005).

We would like, therefore, to ensure that our approach is not restricted to a single climate model, but may be applied to any of the existing ones. To do so, we additionally check whether the method performs equally well for the 200 different sets of parameters obtained by combining each of the 20 AOGCM settings with each of the 10 carbon cycle settings emulated with MAGICC. Figure 1.12 shows the heat map of the magnitude of RMSE in each of these 200 cases.

We find that the average RMSE across all settings is only about 0.04℃, and that in most of the 200 model combinations in MAGICC the average RMSE does not exceed 0.07℃. In general, we find that all combinations produce low approximation errors.[9] The insights from Figure 1.12 could be useful for improving emulation exercises in the future, and for estimating the certainty levels of models' predictions. Overall, we conclude that our emulation technique and the recommended low-dimensional model perform very well for

[9]There are some notable differences among the individual climate and carbon cycle models, such as "Hadley" and "IPSAL", which are both known for strong carbon cycle feedbacks.

the plausible settings of the underlying model parameters.

1.4 Conclusion

New emissions scenarios and socioeconomic pathways are constructed on a regular basis. As the work on the next IPCC Assessment Report has commenced, emulation of large and complex climate models will certainly be on the research agenda. The known resource limitations of running large climate models call for efficient emulation techniques. Our study complements the existing emulation literature by addressing the task of designing efficient input scenarios for emulation. We recommend the use of uncorrelated emissions scenarios for an efficient yet accurate approximation of climate models. These uncorrelated scenarios, based on Chebyshev polynomials, display quite unrealistic emissions paths. However, the purpose of using such scenarios is purely technical—namely, to extract as much information as possible from the complex model.

Using the global temperature anomaly as a predicted response variable, we produce an econometric model—a low-dimensional system of mapping emissions to temperature levels for the twenty-first century. Our simulations confirm that the model performs well on conventional scenarios: the precision of approximation stays high under various settings of climate and carbon cycle parameters. The designed system of equations can be directly implemented in the dynamic stochastic general equilibrium models often used in macroeconomics, allowing one to study optimal policies for dealing with global warming under conditions of uncertainty in terms of social decision-making.

Essay 2

Financial Markets and Climate Models: An Empirical Study on Corn Futures

Financial Markets and Climate Models: An Empirical Study on Corn Futures[1]

Alena Miftakhova, University of Zurich

Walter Pohl, Norwegian School of Economics

Abstract

Some economic sectors—particularly agriculture—are sensitive to weather conditions and hence to the accuracy of available forecasts. Unfortunately, even forecasts of the near future made by the most advanced climate models suffer from relatively low accuracy due to imperfect modeling of the climate system with its highly complex and uncertain nature. Participants in the markets for agricultural commodities would therefore benefit from improved prediction systems—the systems that would facilitate better knowledge about the harvest in the coming season and suggest better strategies to be deployed on financial markets. If these markets are efficient, the prices of those commodities sensitive to weather are expected to reflect the best knowledge about the conditions for the following growing seasons.

In this paper we look for evidence that the expectations of the financial markets for the coming growing seasons are superior to those formed purely from publicly available climate forecasts. We analyze the accuracy of the climate forecasts across corn growing areas of the largest producer of corn—the US—and find no evidence of corn futures markets having more information about future climate conditions than that contained in the publicly available forecasts of the multi-model ensemble studied here.

[1]We thank the participants at the seminar Advances in Computational Economics and Finance at the University of Zurich (2017) and at the Fourth Young Finance Scholars Conference at the University of Sussex (2017) for helpful comments. We are indebted to Per Östberg for his helpful advice and to Dave Brooks for excellent editorial support. Alena Miftakhova gratefully acknowledges financial support from the Swiss National Science Foundation.

2.1 Introduction

The forecasts of state-of-the-art climate models are in great demand with both private and public decision-makers (National Research Council, 2010). Accurate predictions of future changes to the climate facilitate timely actions in agriculture, fishing, energy, and water management among other economic sectors. However, even the short-term forecasts of the most advanced climate models exhibit systematic biases when verified against observational data. This issue stems from the low predictability of the highly uncertain and complex climate system on the time scale of weeks, let alone years or decades.[2]

Decision-makers on all levels—from farmers to insurance companies and governments—are sensitive to the quality of climate forecasts; any information that adds accuracy to these forecasts could help them hedge against potential losses due to adverse climate conditions, or profit should these conditions improve. If agents had access to better prediction systems (e.g., statistical models, heuristics, or climate models with higher skills), they would incorporate this knowledge into the decisions they make with regard to the market. Efficient financial markets would therefore reflect all the intrinsic knowledge of the involved parties in the prices of commodities affected by climate conditions.

In the agricultural sector, the expected weather conditions over the crop growing season are the determinants of planting, harvesting, and storing decisions. In the financial framework, these expectations with regard to the supply of crops form the prices on the corresponding futures markets.

In this study we investigate whether financial markets include any information additional to the predictions supplied by climate models. Our research is inspired by Roll (1984), who finds some predictive power of futures markets for frozen orange juice with respect to corresponding future weather conditions. The high sensitivity of orange plants to freezing temperatures makes the weather the major factor influencing the price; and the participants of the futures market happen to hold expectations superior to those based solely on the forecasts of the US Weather Service.

Today, warmer temperatures might constitute a larger concern than freezes. Extremely hot temperatures that were hardly encountered three decades ago are much more likely to be observed now. In particular, according to Hansen et al. (2012), a large share of the recent temperature anomalies fall more than 3 standard deviations away from what would be the mean of the normally distributed temperatures in the reference period preceding

[2]Climate forecasts are different from weather forecasts in that they include many more processes that characterize the climate system over longer time scales. Weather forecasts, in contrast, cover only up to few days and are derived from the initial state of the atmosphere alone, assuming many other forces remain unchanged over such short periods. The skill (i.e., accuracy) of climate forecasts, therefore, is typically much lower than that of weather forecasts (Kennedy and Kopp, 2000).

rapid global warming.

The ongoing warming of the climate can be beneficial for crops, but the positive effect diminishes rapidly after a certain temperature threshold. For example, for corn and soybeans Schlenker and Roberts (2006) set such a threshold at 30℃ and reveal a strong negative effect of temperatures that exceed this level. Lobell et al. (2014) attempt to quantify the associated damage and attribute a decline of almost 4 percent in the global production of corn to regional changes in weather patterns. In their study, which covers the global weather patterns over the period 1980–2008, only the United States remains an exception to the general warming trend. Hansen et al. (2012) come to a similar conclusion; however, the updated online version of their study includes observations for the last decade and suggests that this pattern—of higher mean temperatures together with wider variation around the mean—does apply to the United States in the summer season.

The forecasts of modern climate models are, in general, to a large extent consistent with the observed trends in temperatures. When considered at a higher resolution, however, they can display strong spatial and seasonal patterns, together with low forecast skills beyond the shortest forecast period of two to four weeks (Slater et al., 2016).

Corn futures markets, in turn, are known to exhibit short-term inefficiencies but are considered efficient in the long run (McKenzie and Holt, 2002). In an early paper Lukac et al. (1988) find evidence of short-term inefficiency by examining 12 trading systems, five of which generate significant profits when trading corn futures. Garbade and Silber (1983) show that corn futures market plays a bigger role in establishing prices than the corresponding cash market—that is, the futures market reflects new information before the cash market does. In their study on whether the corn futures market anticipates weather conditions, Zulauf and Irwin (1998) conclude that the market incorporates some essential information on growing-season droughts into its estimation of the harvest price and is therefore an unbiased predictor of future spot prices.

According to Sumner and Mueller (1989), to have an effect on markets, relevant information (in our case publicly available forecasts and reports) should concern market participants, be issued in a time of decision-making, contain news, and be trusted by the participants (agents). If these criteria are fulfilled, agents adjust their expectations accordingly. For agricultural markets in the US, US Department of Agriculture (USDA) reports are conventionally considered the most relevant and trusted source of information; their public forecasts of corn harvests prove more accurate than those of private agencies (Egelkraut et al., 2003; Good and Irwin, 2006).

The evidence of market reaction to USDA reports is controversial. Some studies find no evidence of any price adjustment on the days reports are issued (Fortenbery and

Sumner, 1993; McNew and Epinosa, 1994). Others report significant reactions of corn futures market to the release of weekly USDA reports (Sumner and Mueller, 1989) and state that the market continues to react even when private forecasts are readily available (McKenzie, 2008). Early research also detected some lasting effect of the days following announcement days, suggesting market inefficiency (Sumner and Mueller, 1989). In later studies, however, market reaction is rapid and adequate, which is consistent with the efficient market hypothesis (Lehecka, 2014). The literature identifies several reports that affect the expectations of traders: the monthly *Crop Forecast* report (Sumner and Mueller, 1989) and those weekly *Crop Production* and *Export Inspections* reports (Colling et al., 1996) issued in the quarter from December to February (Fackler, 1985; Lehecka, 2014), and the monthly *World Agricultural Supply and Demand Estimates* (Isengildina-Massa et al., 2008).

While these monthly reports are not based on climate predictions (Good and Irwin, 2006), the weekly reports are prepared in collaboration with the National Oceanic and Atmospheric Administration (NOAA) and the National Weather Service (NWS)/Climate Prediction Center (CPC),[3] which suggests an implicit link between the forecasts provided by the climate models developed in the CPC and markets' expectations.

In this paper we investigate—using the example of corn (maize), a vital crop—whether a futures market reflects any knowledge additional to the projections[4] of climate models. In particular, we study if the biases in short-term climate forecasts for agricultural regions growing corn correlate with the returns on the futures market. We focus our analysis on the United States, the world's largest corn producer and exporter, in the period 1990–2010.

The rest of the paper is structured as follows. Section 2.2 provides detailed descriptions of the climatic, agricultural, and financial data used in this study and of the way these data are aggregated and aligned together. Section 2.3 outlines the regression methods used in the analysis and presents the obtained results. Section 2.4 discusses the results and the reasons why the relation we seek is hardly identifiable in the chosen framework. Section 2.5 concludes.

2.2 Data collection and processing

The data that can help us discover the abovementioned relation can be divided into three groups. First, the forecasts of climate models offer their predictions of temperature over the region of our interest. The accuracy of these forecasts can be verified against obser-

[3] `https://www.usda.gov/oce/weather/partners.htm`.

[4] We adapt the language of the climate modeling literature and the USDA reports and use the terms "project", "predict", and "forecast" interchangeably.

vational data; any patterns in the forecast errors revealed by such verification represent potential room for improvement. Second, the data on corn production across the country offer a selection of the regions where the crop in question (in our case, corn) grows and where changes to the "normal" climate conditions can significantly affect production volumes. Third, financial data from commodities markets reflect the expectations of the involved market participants with regard to adverse or beneficial developments in the climate conditions. The rest of this section describes in detail all three data groups used in the analysis.

2.2.1 Climate forecasts and observations

As climate data, we utilize the North American Multi-Model Ensemble (NMME) data set (Kirtman et al., 2014). The NMME is a multi-model prediction experiment supported by the National Oceanic and Atmospheric Administration (NOAA), National Science Foundation (NSF), National Aeronautics and Space Administration (NASA), and US Department of Energy (DOE). The database includes the forecasts generated by a set of climate models issuing from the leading US and Canadian climate centers. It also contains their hindcasts. Hindcasts are forecasts initiated at past points in time, for which observations are already recorded. These "forecasts for the past" are useful for analyzing forecast skills and identifying biases and error patterns.

The data are archived and stored at the International Research Institute for Climate and Society (IRI).[5] We use all models of the ensemble that can be pooled together without introducing an obvious bias; that is, we select one model of each type. For example, the two FLOR models—FLOR-B01 and FLOR-A06—are almost identical, with two components improved in the former (Vecchi et al., 2014); we thus use FLOR-B01 here. Similarly, the ECHAM4p5-AC (anomaly coupled) model, not the ECHAM4p5-DC (direct coupled) model, is chosen because anomaly coupling usually corrects the model to match the climatology better.[6] We also exclude the CFSv1 model because the updated CFSv2 performs much better (Saha et al., 2014). Next, the CESM1 model is based on the CCSM4 structure, with enhanced representation of the atmospheric chemistry (Meehl et al., 2013), and therefore is used here. Out of the NASA models—NASA-GEOSS2S, NASA-GMAO, and GMAO-062012—the last one is used. The selected models are listed in Table 2.1.

Typical hindcast periods for the models that participate in the NMME are 1981–2010 and 1982–2010. Every month in this period serves as a starting point for forecasts from 1 to 12 months ahead; these forecasts are called "leads". All projections are made at the resolution 1° latitude by 1° longitude. Out of the set of variables reported by every model in the ensemble we select the variable *tref*, which stands for the monthly mean surface

[5] `http://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/`.

[6] See for example Kirtman et al. (2002) for a comparison of the two coupling strategies.

Table 2.1: Models from the NMME used in this study, with the characteristics of their runs: the period of hindcast, number of runs for each forecast, and the number of leads (months) that the forecasts project.

Model	**Agency**	**Hindcast period**	**Runs**	**Leads**	**Reference**
CanCM4	CMC2	1981–2010	10	0–11	Merryfield et al. (2013)
FLOR	GFDL	1981–2010	12	0–11	Vecchi et al. (2014)
ECHAM	IRI	1982–2010	12	0–7	DeWitt (2005)
GMAO	NASA	1981–2010	12	0–8	Vernieres et al. (2012)
CESM1	NCAR	1980–2010	10	0–11	Marsh et al. (2013)
CFSv2	NCEP	1982–2010	24	0–9	Saha et al. (2014)

temperature level at 2 meters above ground level, measured in degrees Kelvin.[7] This variable has six dimensions—*realization*, meaning model run; *forecast period*, meaning the lead of forecast; *height* (always equal to $2m$ by the definition of the variable); *months since 1960-01-01*, indicating the month of the forecast release; *latitude*; and *longitude*.

The actual observational data are provided by the Global Historical Climatology Network and Climate Anomaly Monitoring System (GHCN + CAMS; Fan and van den Dool, 2008) as a part of the same data set. Adjusted to have the same resolution, the data are stored in the IRI archive and include the same variable (monthly mean surface temperature level) for our period of interest.

A multi-model ensemble is a more reliable and stable forecasting system than any single model that it incorporates (given no "absolutely poor" model in the set—Hagedorn et al., 2005). Studies document a beneficial effect of multi-model pooling on the quality of both average predictions and probabilistic distributions of climate variables (Richardson, 2001; Weigel et al., 2008). The higher forecast skill of a combination of individual predictions of *independent* models stems simply from error cancellation and reductions in the overconfidence[8] of single models (Fritsch et al., 2000; Tebaldi and Knutti, 2007b). More sophisticated methods of combining individual models into an ensemble can sometimes outperform simple averaging (Thompsom, 1976; Krishnamurti et al., 1999; Pavan and Doblas-Reyes, 2000; Rajagopalan et al., 2002); in many cases, however, the equally weighted multi-model mean proves viable if not superior to other weighting schemes (Peng et al., 2002; Kharin and Zwiers, 2002).

[7] We consider temperature and do not consider precipitation because the latter is in general much harder to predict and because the two are highly correlated (Huang and van den Dool, 1993). With regard to anomalies, climate models tend to have similar skills in predicting high temperature extremes and low precipitation extremes, although the performance for precipitation varies greatly among the models (Krakauer, 2017; Becker et al., 2013).

[8] "Overconfidence", in the climate modeling literature, means that the forecasts of a given model have smaller variability than the corresponding observations.

Consistent with this evidence, in the case of the NMME the ensemble proves to be a superior and more stable forecasting system than any of the models taken separately (Zhang et al., 2011; Becker et al., 2014; DelSole and Tippett, 2014; Kirtman et al., 2014). This property applies in particular to our variable of interest, the surface temperature (Krakauer, 2017), and to the geographical region of interest, the continental US in general and its individual regions (Slater et al., 2016). To take advantage of this property, we aggregate the predictions by computing the grand average—that is, we compute the means across all runs of each model and take an average across these means.

Comprising the most advanced climate models, the NMME undergoes regular assessments of the accuracy of the produced forecasts; as results of these assessments, a few weaknesses are currently attributed to the system. The NMME slightly underestimates the frequency of hot extremes and overestimates that of cold extremes (Kirtman et al., 2014). The ensemble is overconfident, with individual participating models exhibiting the same property to an even higher degree. When compared to the historical climatology, the NMME predicts both the mean and percentiles of the temperature distribution better, but its performance becomes rather poor beyond the shortest leads (Slater et al., 2016; Krakauer, 2017). The same studies document strong spatial and temporal patterns in the actual forecast skill, with particularly low (almost none) potential skill in the Midwest region—our major region of interest, where most of US corn production is concentrated.

2.2.2 Corn production

The United States has traditionally been the leading producer and leading exporter of corn. According to the Foreign Agricultural Service (FAS) of USDA, for the last two decades the United States has maintained a proportion of around 40 percent of global corn production, as shown in the top panel of Figure 2.1. With the latest statistic of over 35 million hectares reserved for this crop, production has grown from 8 billion bushels (201 million tonnes) in 1990 to 15 billion bushels (385 million tonnes) in 2016. Most of the corn produced is used for ethanol production or livestock feed; with respect to the latter, corn covers almost the entire volume (96 percent) of feed grains used in the country.[9] Around 14 percent of US corn production is exported to other countries,[10] with the largest demand coming from Japan, Taiwan, and Mexico. The global market for corn chiefly supplies the market for livestock and poultry feed; the traditional substituting goods for corn on this market are grain sorghum, barley, oats, and low-quality rice and wheat (Wisner and Baldwin, 2004).

Most of the country's corn production is concentrated in the Midwest region, in the so-called Corn Belt, which has the optimal climate conditions for corn plants. The crop grows

[9] USDA, `https://www.ers.usda.gov/topics/crops/corn-and-other-feedgrains/background/`.
[10] US Grains Council, `https://grains.org/buying-selling/corn/`.

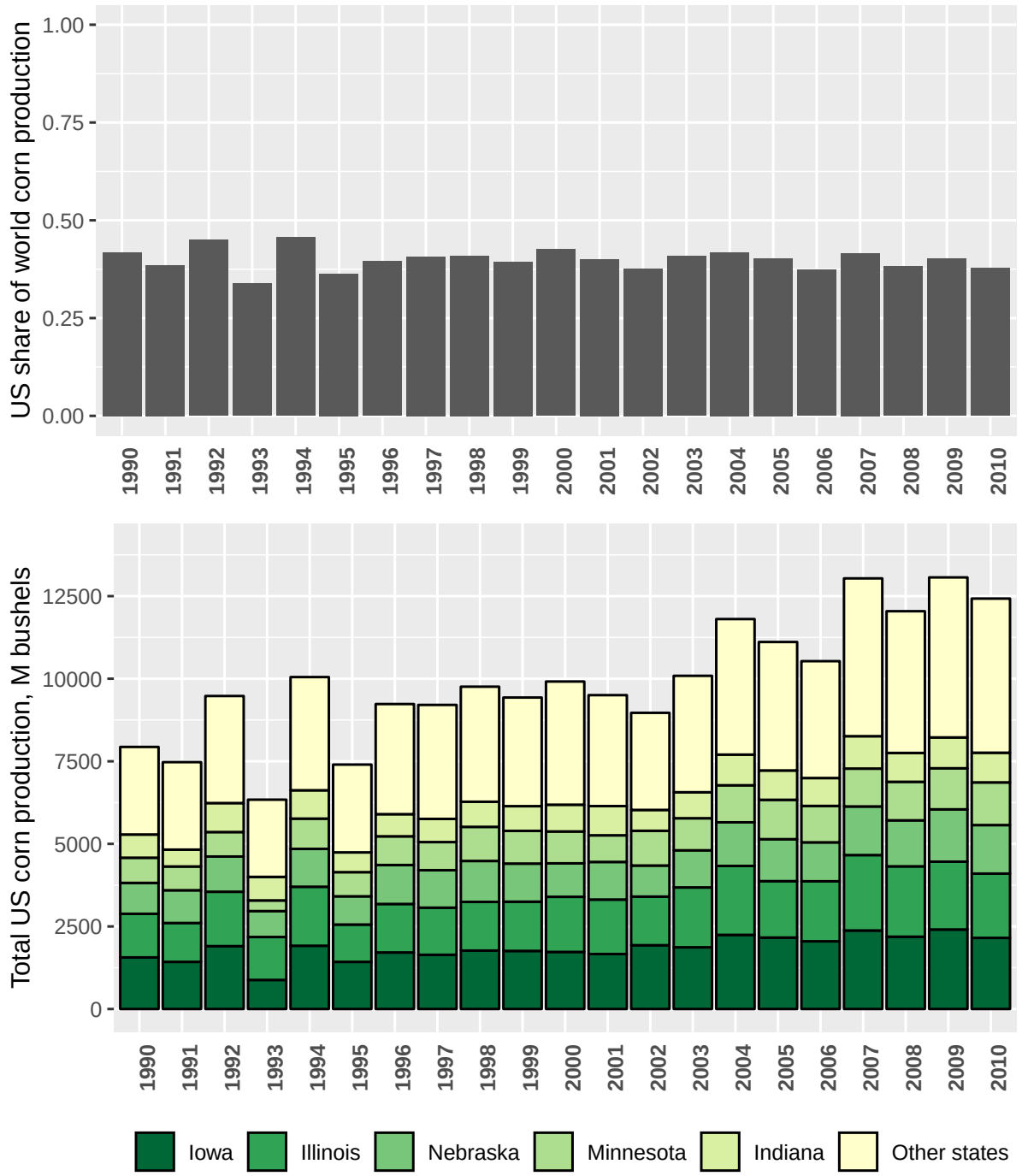


Figure 2.1: The share of the United States in global corn production over the sample period (top) and the distribution of US corn production among the leading five states and the rest of the country (bottom).
Source: Reproduced from the FAOSTAT database, Food and Agriculture Organization of the United Nations, `http://www.fao.org`.

in almost every state; the leading states traditionally have been Iowa, Illinois, Nebraska, Minnesota, and Indiana—together they cover over 60 percent of the total production volume (see the bottom panel of Figure 2.1).

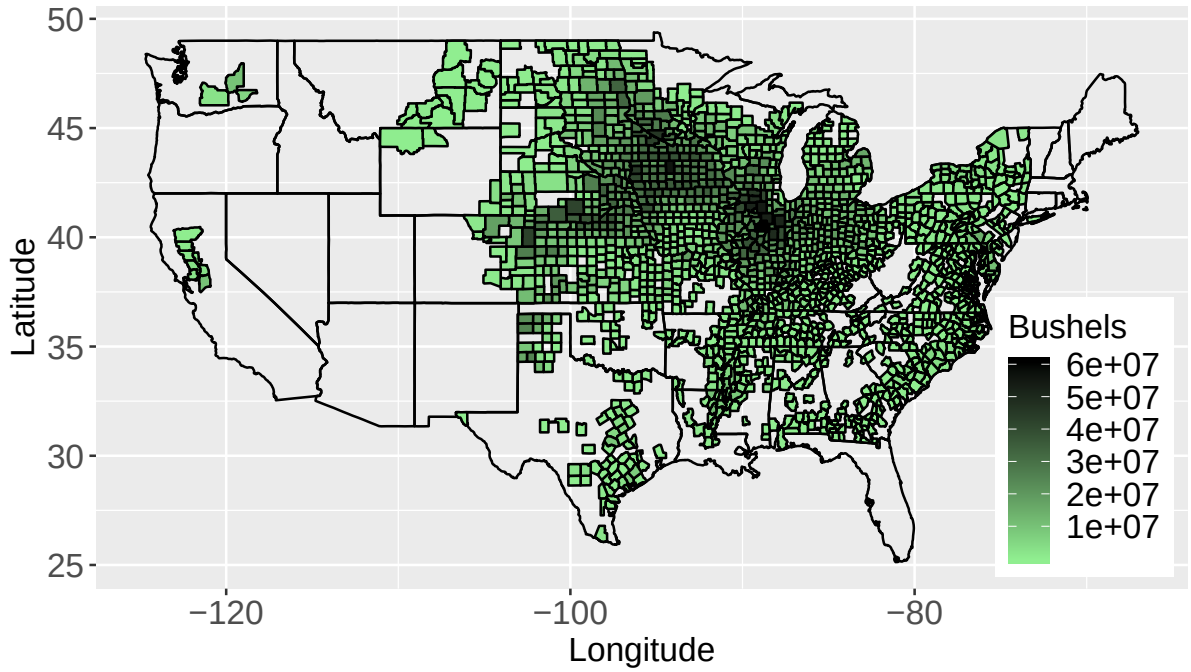


Figure 2.2: Corn production in the United States in 2010 in bushels, by counties. *Source: Reproduced from the data in the Quick Stats database of NASS, USDA.*

The county-level data for this study are obtained from the Quick Stats database of the USDA National Agricultural Statistics Service (NASS).[11] Figure 2.2 shows an example—the distribution of the production of corn by volume (in bushels) in 2010. Using the county-level data, we can later select these areas where the crop is grown from the countrywide climate forecasts and weight counties by their contribution to the country's total production of corn.

2.2.3 Futures on corn

In the US over a million farmers sell corn to operators of grain elevators, to other grain processors, or to subsequent producers of livestock and poultry. Buyers and resellers, in turn, hedge their risks related to the supply of corn on the futures market. The most important factors that cause price shocks on this futures market are weather conditions, internal and international policies, and fluctuations in other countries' production (Wisner and Baldwin, 2004). The futures market constitutes not only a hedging platform, but also the major indicator of price movements, which subsequently transfer to the local cash markets. Farmers, buyers, and other market participants therefore have to monitor this market continuously.

The dominant market for corn futures is the Chicago Board of Trade (CBOT)—the leading futures and futures-options exchange, where the prices for the commodity form.

[11] The data base is available at `https://quickstats.nass.usda.gov/`.

Table 2.2: Summary statistics for the first 10 traded corn futures contracts; monthly data from January 1990 to December 2010. Standard deviations are reported in brackets.

	N obs.	Avg. price, USD	Avg. volume	Avg. return
Contract 1	252	1,647.315	1,058,597	−0.004
		(1,073.571)	(792,041)	(0.072)
Contract 2	252	990.523	604,394	−0.002
		(516.084)	(629,614)	(0.071)
Contract 3	252	668.331	243,189	−0.001
		(286.250)	(332,861)	(0.067)
Contract 4	252	622.688	119,125	−0.0001
		(220.445)	(201,149)	(0.062)
Contract 5	252	554.102	62,127	0.001
		(142.422)	(115,017)	(0.056)
Contract 6	252	448.588	37,365	0.002
		(92.708)	(100,419)	(0.049)
Contract 7	243	307.767	13,020	0.003
		(74.448)	(38,539)	(0.043)
Contract 8	179	269.208	9,101	0.005
		(96.371)	(25,706)	(0.045)
Contract 9	147	336.549	5,955	0.003
		(87.573)	(16,628)	(0.043)
Contract 10	130	282.979	3,859	0.013
		(98.068)	(9,158)	(0.040)

Note: The maximum number of observations is 252, which stands for the 240 months in the sample period plus the 11 following months, for which forecasts are issued in the last month of the sample, and one month preceding the first month of the sample.

There are five corn futures contracts scheduled for each year—with deliveries in March, May, July, September, and December; these contracts trade in the unit of 5,000 bushels. December is the new-crop contract, whereas the other four contracts trade the harvest of the preceding year.

We use the monthly data on futures prices and returns because at this frequency the data reflects the overall expectations of the market participants for the coming growing season. Intraday fluctuations, day-of-the-week effects, and price shocks driven by news releases or by other factors are therefore not subjects of concern. All the financial data are retrieved from the *Bloomberg* data base, Bloomberg Finance L.P.[12] Because of many missing values and negligible volumes of trade reported in the database prior to the 1990s, we start the sample period from 1990, which, given the length of the hindcasts from the NMME, yields a sample of 240 months from January 1990 to December 2010. Table 2.2 shows the statistics for monthly data on corn futures over the examined period 1990–2010 for the first 10 traded futures contracts.

[12]All data on corn futures were retrieved via the Bloomberg Terminal in May 2016.

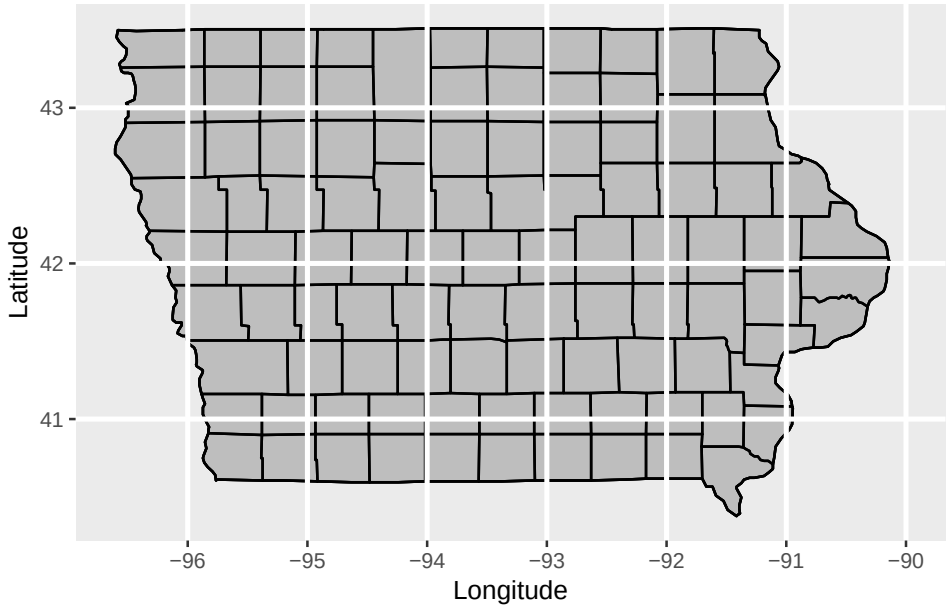


Figure 2.3: An example of a counties map superimposed on the corresponding temperature map, for Iowa state.

Within the sample period the nearest futures contract was, unsurprisingly, traded the most and at the highest price, whereas the contracts further in the future were much less liquid. The large standard deviations of contracts' prices come mostly from substantial gradual reduction in the prices over the sample period. For the first contract, for example, the settlement prices exceeded USD 3,000 in 1990 and shrank to less than USD 500 by 2010. The returns are on average indistinguishable from zero and more volatile for nearer contracts. Our financial variable of interest is monthly return on futures contracts calculated as the month-to-month *relative* change in the settlement price.

2.2.4 Merging the temperature data with the yields data

The resolution of the climate forecasts in our database is one by one degree of latitude and longitude, whereas the production of corn is reported on the finer level of counties. To merge the two data sets, we superimpose the temperature grid on the map of the counties, as demonstrated in Figure 2.3 with the example of Iowa. This overlap divides each county into subregions, depending on how a county is cut by the grid lines.

This division enables us to distinguish between the resulting subregions both by the forecasted/observed temperature values and by the importance of each subregion in the total volume of corn production (or, equivalently, by the share of the county containing

Table 2.3: An example of areas computed for counties and the corresponding temperature values.

	Area, m	State	County	Temperature, K
1	10,807,960	ALABAMA	AUTAUGA	285.232
2	32,114,032	ALABAMA	AUTAUGA	284.975
3	737,726,630	ALABAMA	AUTAUGA	286.342
4	708,720,993	ALABAMA	AUTAUGA	286.350
5	1,038,387,369	ALABAMA	BULLOCK	286.350
6	576,290,247	ALABAMA	BULLOCK	286.359
7	2,432,907,880	ALABAMA	DALLAS	286.342
8	323,714,283	ALABAMA	ELMORE	284.975
9	1,355,337,930	ALABAMA	ELMORE	286.350
10	1,426,284,024	ALABAMA	GREENE	285.410
11	269,651,722	ALABAMA	GREENE	286.354
12	562,754,188	ALABAMA	HALE	285.410
13	612,556,137	ALABAMA	HALE	285.232
14	303,434,019	ALABAMA	HALE	286.354
15	262,303,242	ALABAMA	HALE	286.342

any given subregion in the entire country's production).

Next, we would like to calculate the area-based weights for each such subregion within a county. The default projection of the maps from the globe to a flat surface, however, does not respect distances and would distort the calculation. Therefore, we first *re-project* both the temperature maps and the counties' maps onto a new coordinate reference system (CRS) that preserves areas. Then, the areas can be calculated for each temperature level in each county.[13] Table 2.3 shows an example of the resulting calculations of areas and the assignment of temperatures for each subregion, using six counties in Alabama.

2.2.5 Corn growth and development

Favorable temperatures are the major factor for the growth of the corn plant. Growth might begin at a temperature as low as 8℃ (281.15 K), but the appearance and growth of new leaves is conditional on plants' exposure to much higher temperatures, usually 22–32℃ (295.15–305.15 K). Temperatures exceeding the threshold of 32℃ turn out to be adverse for the plant. In fact, the so-called *stress index* for corn is derived from the number of days when the temperature stays above 32℃ (Westgate et al., 2004).

The most critical period for the plants' growth—the period that determines grain yields—

[13]The initial coordinate reference system is "proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0". We re-project it to a new CRS, "+proj=aea +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0", as recommended in Kennedy and Kopp (2000).

is from shortly before to shortly after silking,[14] when the ears of the plants develop. During this stage, the ears emerge and grow to their full length, whereby the number of kernels per ear is also determined. After the kernels form, both the rate of their development and their potential to reach their maximum weight are highly sensitive to temperatures. Too high temperatures, for example, can accelerate kernel growth and thereby shorten the period during which the kernels grow, leading overall to lower yields (Hicks and Thomison, 2004).

Water deficiency can also affect the plants' growth adversely, but the effect is not as apparent: water-deficient plants tend to have smaller leaves, which reduces photosynthesis. At the later stage of the kernel development, lack of water can lead to lower kernel mass. Yet the effect can be offset by a higher photosynthesis rate per unit of leaf area—a rate that is not affected by the water deficit.

The growing season of corn plants is specific to regional agricultural and climate conditions. In the southern states planting might start as early as mid-February, while for most of the Corn Belt areas the optimal planting period is from late April to early May. In the central Corn Belt actual plant growth starts in early May and lasts until mid-July, followed by the reproductive stage of ear development. Harvesting starts in late August and continues until early November. Air temperature is an important factor in this late stage as well: the rate of grain drying for later storage depends on the number of sequential warm days. Most of the harvest is stored locally until its delivery in later months, in December at the earliest (Maier, 2004).

2.2.6 Relating climate forecasts and futures contracts

Combining all three data sets involves, first, selecting the climate predictions that correspond to the corn growing season and, second, carefully matching them to the deliveries scheduled by futures contracts.

If we consider the growing season of year t, the corresponding climate forecasts would span the period from $March_t$ to $October_t$. The first such forecast, for $March_t$, is released in $April_{t-1}$ and is an 11-months-ahead forecast. In the next month, May_{t-1}, two forecasts—for $March_t$ and $April_t$ (for 10 and 11 months ahead, respectively)—become available. The set of available forecasts for the growing season t thereby grows until $October_{t-1}$. The forecasts issued in the following five months (from $November_{t-1}$ to $March_t$) cover the entire growing season of the subsequent year—from $March_t$ to $October_t$. Starting from $April_t$ the set reduces by $March_t$, and shrinks further as the months pass, until the last forecast issued in $October_t$ for the month itself. Every forecast in the set selected for a

[14] Silking is the stage of the plants' growth when silk like strands emerge from the ends of plants' ears (USDA Crop Progress, Terms and Definitions; `https://www.nass.usda.gov/Publications/National_Crop_Progress/terms_definitions.php`).

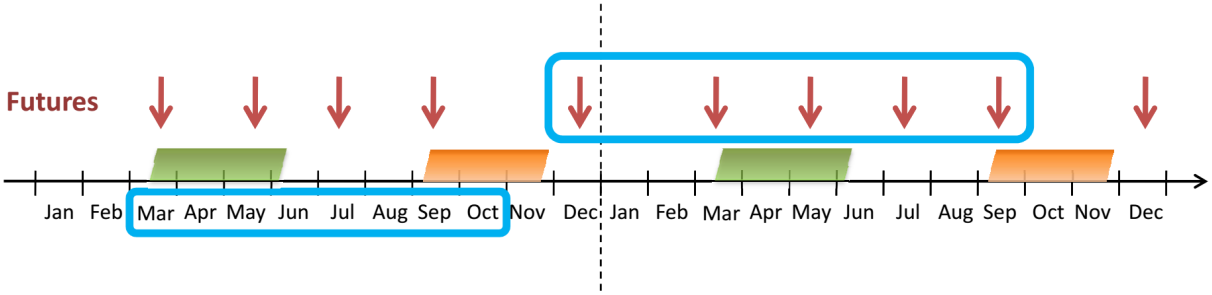


Figure 2.4: Time line relating the corn growing season and futures contracts. The time line covers two years. Planting seasons are marked green; harvesting seasons in orange. The delivery dates of futures contracts are marked with arrows. The growing season of the first year and the futures contracts for the corn grown in this season are highlighted with blue rectangles.

particular growing season is thus characterized by the predicted (forecast) month and the month of issue. The distance between the two determines the *lead*, or how far ahead the forecast predicts.

Because harvesting starts in September, the nearby December futures contract is the first one to trade the freshly collected corn and is usually referred to as the "new crop contract". The full set of futures contracts related to the harvest of year t, therefore, is $\{December_t, March_{t+1}, May_{t+1}, July_{t+1}, September_{t+1}\}$. For illustration, Figure 2.4 depicts a time line of two years with the growing and harvesting seasons for corn marked in green and orange, correspondingly. The scheduled deliveries for the standard corn futures contracts are shown with arrows. The growing season of the first year and the set of the futures contracts that trade the harvest of that year are marked with blue rectangles.

Each climate forecast for a growing season in a particular year is therefore potentially related to the five corresponding futures contracts. For example, for a forecast for $March_t$ (the month of forecast) released in $January_t$ (the month of issue), the $5th$, $6th$, $7th$, $8th$, and $9th$ traded futures contracts can be used as predicting variables, while the nearest four futures contracts trade the harvest of the previous year.

2.2.7 Data aggregation

The futures market establishes one, universal price for the standardized contract on corn. All spatial data therefore have to be aggregated into one index that can be related to the price in any given period. Essentially, for both the climate forecasts and realizations (observations), we take the weighted averages across the crop growing areas—weighted by the subregions' area and counties' production in a given year, as outlined in Section 2.2.4.

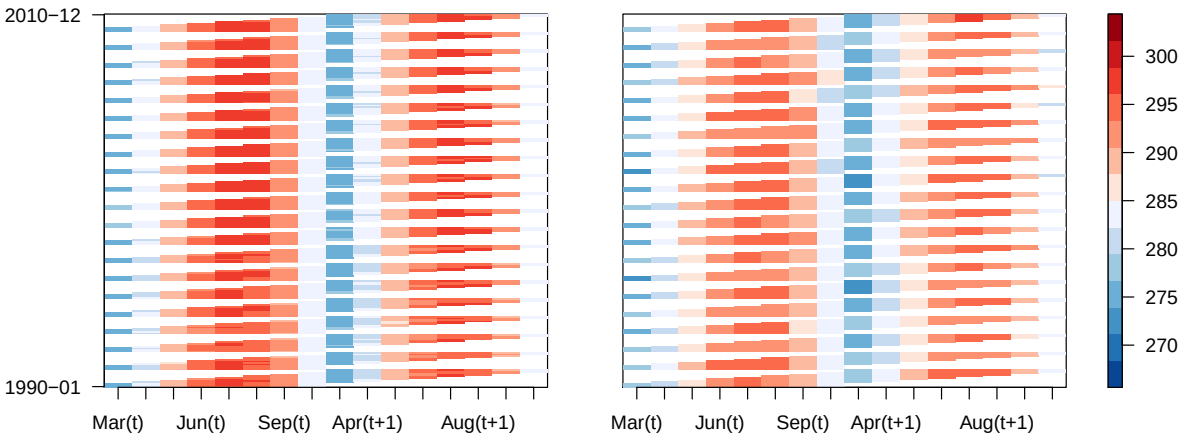


Figure 2.5: The forecasted (left) and real (right) temperature indices, $Tfind$ and $Trind$, plotted by the months of forecast. The vertical axis spans the sample period—from January 1990 at the bottom to December 2010 at the top. The color scale indicates the temperature levels, in degrees Kelvin.

For each month k of forecast issue and each forecast lead l, the **temperature forecast index**, $Tfind_{k,l}$, is calculated as a sum of the forecasted temperature levels, $Tf_{i,j,k,l}$, across all temperature values j and counties' subregions i, weighted both by the area weights, w_i^A, and by the production weights, $w_{i,n}^P$, in every year n of the forecast,

$$Tfind_{k,l} = \sum^{i} \sum^{j} w_i^A w_{i,n}^P Tf_{i,j,k,l}. \tag{2.1}$$

The production weight $w_{i,n}^P$ for each subregion i is proportional to its county's share in the total corn production in year n; the area weight, w_i^A, for each subregion i is proportional to the share of the subregion in the county's area.

The **real temperature index**, $Trind_m$, for every month m is constructed analogously from the observed temperature levels, $Tr_{i,j,m}$,

$$Trind_m = \sum^{i} \sum^{j} w_i^A w_{i,n}^P Tr_{i,j,m}. \tag{2.2}$$

Figure 2.5 juxtaposes the temperature forecast index with the real temperature index, both plotted against the forecasted months, for the entire sample period.

The deviation of the forecasts from the observational data—the **forecast error index**, $Tratio_m$—is represented by the ratio of the temperature forecast index of a month k for a lead l to the real temperature index in the forecasted month $k + l$,

$$Tratio_m = \frac{Tfind_{k,l}}{Trind_m}, \ m = k + l. \tag{2.3}$$

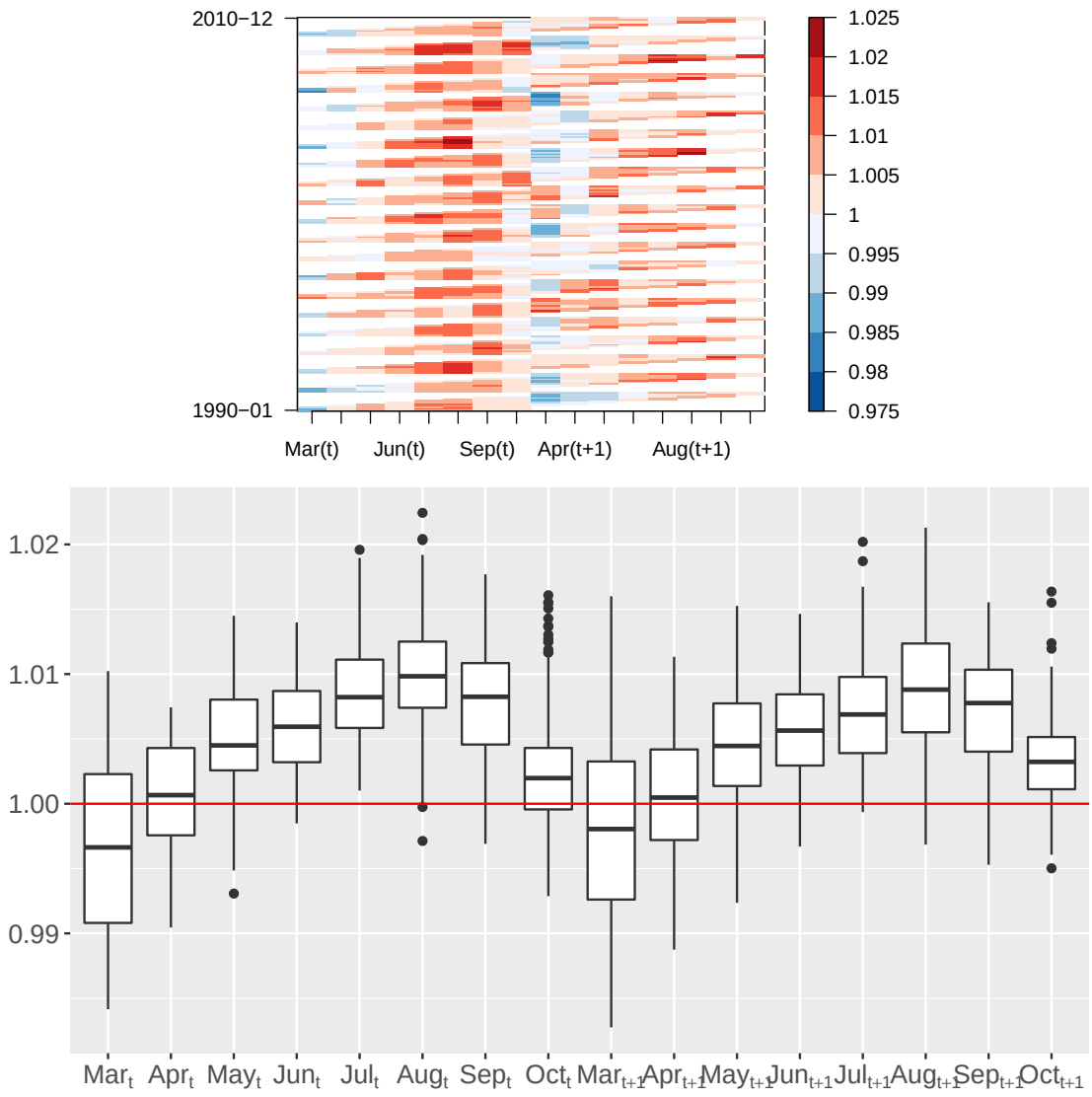


Figure 2.6: *Top:* The forecast error index, *Tratio*, plotted by the months of forecast (horizontal axis) and the months of issue (vertical axis), analogous to Figure 2.5. The vertical color scale indicates the value of the ratio in (2.3). *Bottom:* Box plots of *Tratio* index plotted by months of forecast.

Figure 2.6 plots our main variable of interest—the forecast error index—and demonstrates that on average for most of the months of forecast climate models predict too high temperatures; the biases increase towards the peaks of growing seasons and are of comparable sizes for the closer and the further growing seasons.

For every month of issue and every month of forecast, we collect all the data (the forecast index, the real temperature index, the forecast error index, and the futures returns) into a summary table; part of this table is reproduced in Table 2.4 and includes the returns for the first three out of the five associated futures contracts. The prominent features of the data set in question, such as repetitive values in multiple columns, are addressed in the next section.

Table 2.4: An example of the data for the first months of issue—the months of forecast, temperature forecast index, real temperature index, forecast error index, and returns for the first three futures contracts.

Issue m.	Forecast m.	Tfind	Trind	Tratio	Contr.1	Contr.2	Contr.3
1990-01-01	1990-03-01	275.248	277.892	0.990	−0.015	−0.015	−0.019
1990-01-01	1990-04-01	282.795	282.325	1.002	−0.015	−0.015	−0.019
1990-01-01	1990-05-01	289.394	286.795	1.009	−0.015	−0.015	−0.019
1990-01-01	1990-06-01	295.126	294.098	1.003	−0.015	−0.015	−0.019
1990-01-01	1990-07-01	297.950	295.197	1.009	−0.015	−0.015	−0.019
1990-01-01	1990-08-01	297.335	294.683	1.009	−0.015	−0.015	−0.019
1990-01-01	1990-09-01	292.199	291.679	1.002	−0.015	−0.015	−0.019
1990-01-01	1990-10-01	283.890	283.566	1.001	−0.015	−0.015	−0.019
1990-02-01	1990-03-01	274.961	277.892	0.989	0.036	0.035	0.024
1990-02-01	1990-04-01	283.112	282.325	1.003	0.036	0.035	0.024
1990-02-01	1990-05-01	289.802	286.795	1.010	0.036	0.035	0.024
1990-02-01	1990-06-01	295.524	294.098	1.005	0.036	0.035	0.024
1990-02-01	1990-07-01	298.454	295.197	1.011	0.036	0.035	0.024
1990-02-01	1990-08-01	297.955	294.683	1.011	0.036	0.035	0.024
1990-02-01	1990-09-01	292.588	291.679	1.003	0.036	0.035	0.024
1990-02-01	1990-10-01	284.324	283.566	1.003	0.036	0.035	0.024
1990-03-01	1990-03-01	276.708	277.892	0.996	0.006	0.004	0.006
1990-03-01	1990-04-01	283.077	282.325	1.003	0.006	0.004	0.006
1990-03-01	1990-05-01	289.804	286.795	1.010	0.006	0.004	0.006
1990-03-01	1990-06-01	295.448	294.098	1.005	0.006	0.004	0.006

2.3 Methods and results

This section presents the results of an econometric analysis of the data. The structure of our data set is rather nontrivial: first, every 8 observations are the forecasts issued in the same month for different leads; second, for any month of forecast there are 12 forecasts issued in the 11 preceding months and in that month itself; finally, in every month of issue, the return on the futures contracts associated with a set of forecasts for the same growing season has the same value. The first part of this section formulates the basic model for estimation and provides the theory and the practical implications of dealing with a data set clustered in multiple ways. The rest of the section presents the results of the estimation and a few ways to refine the analysis.

2.3.1 Simple regression model

In its general form the regression equation relates the climate forecasts for the following growing season to the returns on the futures market,

$$Tratio_m = \beta_0 + \boldsymbol{R'_m}\boldsymbol{\beta} + \varepsilon_m, \tag{2.4}$$

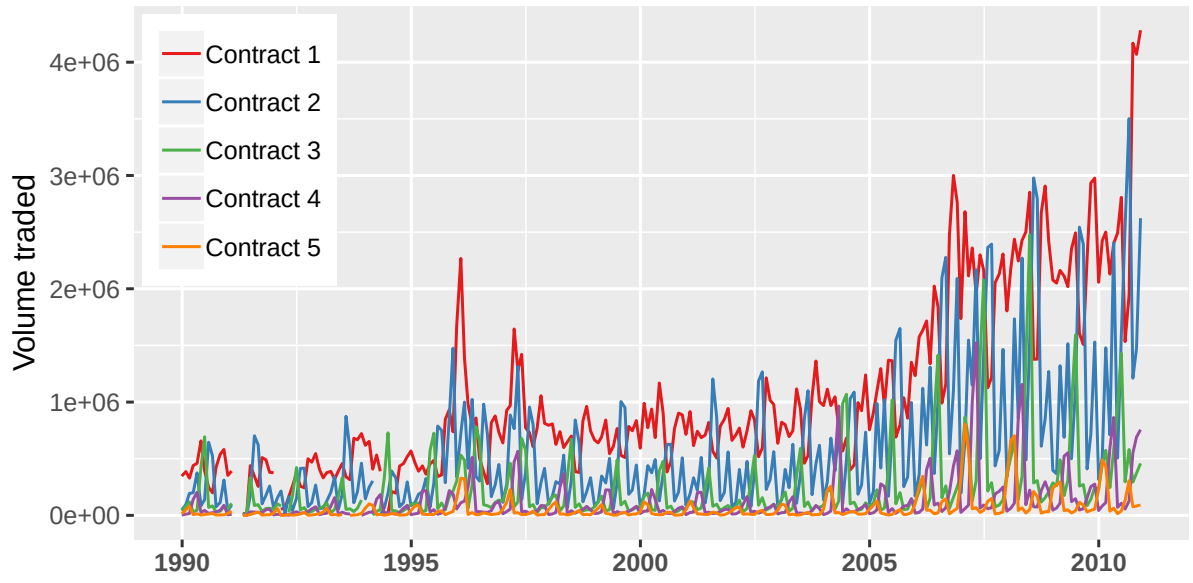


Figure 2.7: Volumes traded for the five futures contracts related to the corresponding growing seasons in the sample period. Contract 1 is the nearest (December) contract; Contract 5 is the contract with the latest delivery.

where, for every forecast month m in the sample period, $Tratio_m$ is the ratio of the forecast index to the real temperature index from Equation (2.3) in Section 2.2.7, $\boldsymbol{R_m}$ is a vector of returns-related covariates, and ε_m is an error term.

As explained in Sections 2.2.6 and 2.2.7 above, each "observation" m is the error of the forecast issued for a certain month in the sample period related to a set of the five futures contracts that trade the harvest of the same growing season. In most of the observed months the first contract in the set of these five contracts—the December contract—covers more than 90 percent of the total volume of trade for the harvest of that season (see Figure 2.7). At the same time, the correlation of the returns on the December contract with returns on the other four contracts is close to 1, and its correlation with the volume-weighted average across the five contracts is indistinguishable from 1. Therefore, we use the return on the new-crop December contract as a single regressor $R1$ in (2.4).

2.3.2 Clustering and correlation in the data

In the simplest setting, the error term in Equation (2.4) can be assumed to be independently identically distributed (*iid*), which greatly simplifies the subsequent statistical inference. However, it is very rarely the case in econometric analysis that this assumption applies to a data set. The errors, instead, might exhibit non-constant variance across the observations (heteroskedasticity), or correlation within some subgroups (clusters) of the data set. Either issue invalidates the statistical inference made under the *iid* assumption.

Estimation in the presence of clustered correlations

The structure of our data set suggests that *clustering* is the form of heteroskedasticity that we are likely to encounter. In particular, the observations come in groups that have the same values of the regressor, $R1$, and might be correlated based on the unobserved effects that a particular futures contract could not capture. The forecasts, in turn, come in groups issued in the same months and can correlate based on the skill of the forecasting system at the time of forecast release. On the other hand, every observed month has 12 forecasts preceding the observation—a group that shares a common feature, the predictability of the weather conditions in that month. Failure to control for these potential within-cluster correlations of the error term can lead to strongly underestimated standard errors, and, consequently, spurious statistical significance (Cameron and Miller, 2015).

The simple regression Equation (2.4) containing only one non-constant covariate can be rewritten in a cluster-specifying form,

$$Tratio_{u,g} = \beta_0 + \beta_1 R1_{u,g} + \varepsilon_{u,g}, \tag{2.5}$$

where g is a group index, $g = \{1, 2, \cdots, G\}$, and u is a within-group index, $u = \{1, 2, \cdots, U\}$, assuming equal size U for all groups.

One conventional way to introduce within-group correlation is to decompose the error term into the group-specific component, μ_g, which is assumed to capture all cluster-specific correlation, and the remaining *iid* component, $\eta_{u,g}$,

$$\varepsilon_{u,g} = \mu_g + \eta_{u,g}. \tag{2.6}$$

Cluster effects are thus one type of *unobserved effect* and can be treated with appropriate econometric tools—according to the assumption of presence or absence of correlation between these effects and the explanatory variable.

In the case of within-cluster correlation of *only* the error term and not the regressor, formally stated as $\mathbb{E}\left[R1'_{u,g}\mu_g = 0\right]$, the traditional OLS estimator is consistent and asymptotically normal. However, a well-established result[15] for a univariate regression analysis of a balanced (containing equally sized groups) data set is that the OLS variance estimate is underestimated by a factor v_g approximately equal to

$$v_g = 1 + \rho_R \rho_\varepsilon (U - 1), \tag{2.7}$$

[15]This result is established by Kloek (1981), Greenwald (1983), and Moulton (1986); see also Cameron and Miller (2015) or Angrist et al. (2013) for an overview of cluster-related bias correction.

where ρ_R is the within-group correlation of the regressor and ρ_ε is the within-group correlation of the error term. The (in most cases) downward bias of the standard errors is therefore the stronger, the larger the groups and the higher the errors' and regressors' correlations within these groups.

After the OLS estimation, the standard errors can be adjusted for the within-cluster correlation, provided the number of clusters is large relative to their sizes (Wooldridge, 2010). The adjusted variance matrix is

$$\widehat{Var}\left[\boldsymbol{\beta}\right] = \left(\sum_{g=1}^{G} \mathbf{X}_g'\mathbf{X}_g\right)^{-1} \left(\sum_{g=1}^{G} \mathbf{X}_g'\hat{\boldsymbol{\varepsilon}}_g\hat{\boldsymbol{\varepsilon}}_g'\mathbf{X}_g\right) \left(\sum_{g=1}^{G} \mathbf{X}_g'\mathbf{X}_g\right)^{-1}, \tag{2.8}$$

where $\mathbf{X}$ is the matrix of covariates (which for the model in (2.5) includes the vectors of the constant and regressor) and $\hat{\boldsymbol{\varepsilon}}$ is the vector of the estimated residuals. Because the intragroup correlation in (2.8) can take arbitrary forms, this variance matrix is robust both to within-cluster correlations and to heteroskedasticity in general.

The OLS estimator is consistent but loses its efficiency when model errors are correlated within clusters. A more efficient estimator is the *random effects* (RE) estimator, which uses the feasible generalized least squares (FGLS) framework to obtain the estimates of coefficients in the model. It requires, however, that the variance matrix be explicitly and correctly specified, and the efficiency gains are relevant only under the assumption of strict exogeneity between μ_g and $R1_{u,g}$ (Wooldridge, 2010).

If the cluster effects are correlated with the regressor, $\mathbb{E}\left[R1_{u,g}'\mu_g \neq 0\right]$, both the OLS and FGLS estimators become inconsistent. An alternative, viable model for estimation is the *fixed effects* (FE) model, which de-means the data group-wise before the OLS estimation of the resulting specification

$$(Tratio_{u,g} - \overline{Tratio}_g) = \beta_1(R1_{u,g} - \overline{R1}_g) + (\eta_{u,g} - \overline{\eta}_g), \tag{2.9}$$

where $\overline{Tratio}_g$, $\overline{R1}_g$, and $\overline{\varepsilon}_g$ are the corresponding group-wise averages of the regressand, regressor, and error term, respectively. The advantage of the fixed effects estimator is that it eliminates the cluster-invariant component, μ_g, and therefore to a large extent controls for cluster-specific correlation. Naturally, all cluster-invariant covariates are eliminated by the transformation (2.9) as well.

In practice it is often the case that the linear component, μ_g, does not fully reflect the within-group correlations. In our data set, for example, introducing linear season-specific effects might still leave some room for month-specific correlation of the forecasts and hence of the model errors. In such cases the idiosyncratic component, $\eta_{u,g}$, violates the *iid* assumption and exhibits some additional within-cluster correlation (and/or heteroskedas-

ticity); the FE model thus needs to be complemented with cluster-robust variance matrix adjustment analogous to (2.8). The resulting variance matrix estimator for the FE model is robust to any kind of intracluster correlation and arbitrary heteroskedasticity, for a large enough number of clusters relative to their sizes.

Testing for cluster effects

Statistical testing offers several options to detect unobserved effects and intragroup correlations in a model. The presence of unobserved effects can be generally tested for using the test developed by Wooldridge (2010)—the test checks if the non-diagonal elements of the estimated covariance matrix are significantly different from zero. The test statistic

$$W = \frac{\sum_{u=1}^{U} \sum_{g=1}^{G-1} \sum_{s=g+1}^{G} \hat{\varepsilon}_{u,g} \hat{\varepsilon}_{u,s}}{\left[\sum_{u=1}^{U} \left(\sum_{g=1}^{G} \sum_{s=g+1}^{G} \hat{\varepsilon}_{u,g} \hat{\varepsilon}_{u,s} \right)^2 \right]^{1/2}} \tag{2.10}$$

under the null hypothesis of no intragroup correlations asymptotically follows the standard normal distribution. The rejection of the null hypothesis indicates the presence of within-cluster correlations and suggests modifying the model, changing the estimator, or adjusting the estimated standard errors.

The key property of correlation between the unobserved group effects, μ_g, and the regressor, $R1_{u,g}$, can be tested for with the Hausman test (Hausman, 1978). The test effectively compares FE and RE estimators, $\hat{\boldsymbol{\beta}}_{FE}$ and $\hat{\boldsymbol{\beta}}_{RE}$, of the model. The test statistic

$$H = \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE} \right)' \left[\widehat{Var} \left(\hat{\boldsymbol{\beta}}_{FE} \right) - \widehat{Var} \left(\hat{\boldsymbol{\beta}}_{RE} \right) \right]^{-1} \left(\hat{\boldsymbol{\beta}}_{FE} - \hat{\boldsymbol{\beta}}_{RE} \right) \tag{2.11}$$

is asymptotically chi-squared distributed. If the difference between the two estimators is statistically significant, the RE estimator is considered inconsistent since it departs from the consistent FE estimator. Because the RE relies on the absence of the correlation, this null hypothesis can be rejected.

Multiple ways to cluster

In each case a decision on a particular way to adjust the estimation procedure is made based on the presumed structure of the correlation and on the nature of the unobserved effects in the data set. There are often multiple ways to cluster the data, with different clusters nested (included into one another; e.g., months and years) or overlapping (e.g., states and years). As for our data set, it can be clustered in the following multiple ways.

Month of issue. Every eight observations in the data set are a set of eight forecasts of different leads issued in a certain month (the column "Issue m." in Table 2.4). The

accuracy of climate forecasts is known to depend strongly on the initial conditions of the individual runs of climate models and on the specification and parameterization of these models (Tebaldi and Knutti, 2007a). The errors, consequently, might be correlated for the forecasts produced by the same model or forecast system at the same time point, given the specified starting points and the state of the forecasting system itself.

Yet in our case we expect the within-month-of-issue correlation to be only a minor concern, mainly for two reasons. First, because our data are hindcast data, the forecast system does not change over time while issuing the forecasts. Second, the sensitivity of the forecasts to the initial conditions is treated by averaging across the model runs—each of the runs has slightly perturbed initial conditions—and across all selected models in the ensemble.

Month of forecast. Similarly, the errors can be correlated for the forecasts that predict the same month (note the repetition of the values in the column "Forecast m." in Table 2.4). To the extent that the factual temperature could not be predicted and was *unexpectedly* warm or cold, the errors will be correlated for the same realization. In Equation (2.5) this property of the error term would refer to the unexpected realization of weather in a particular month, not captured by either the forecast or market expectations.

Because the observations within such clusters are the forecasts for the same month from 11 to 0 months ahead, linear components (i.e., fixed effects) are unlikely to capture the correlation completely. We therefore do not impose a particular structure on the error term but rather adjust the standard errors of our estimates as in (2.8) such that they are robust to within-month-of-forecast correlation.

Values of the regressor. The forecasts issued in the same month can refer to the expectation for the growing season of the same year or of the following year. Therefore, the regressor $R1$ (the column "Contr.1" in Table 2.4) is not clustered strictly by the month of issue but rather by the month of issue and season. Because our only non-constant regressor is this "clustered" regressor, the errors within such clustering are likely to be correlated—capturing the factors that the market could not predict in a particular month.

This clustering, however, does not require a separate adjustment. By construction, the regressor is perfectly correlated within these clusters and hence captures the linear effect μ_g; the potential residual correlation is accounted for when we consider month-of-issue or season clustering, which nest the current clustering and are therefore more conservative.

Growing season. Clustering at the growing season level (equivalent to the year level in our case) is more than simply aggregating the more detailed month-of-forecast clustering. It, as well, reflects the predictivity of the weather conditions—on a more aggregated,

Table 2.5: Test statistics and their statistical significance for the Wooldridge test and Hausman test for the four potential ways of clustering in the data set.

	Test statistic	
	Wooldridge	Hausman
Month of issue	0.5984	4.5714^{**}
Month of forecast	7.147^{***}	0.3077
Regressor $R1$	-14.58^{***}	–
Growing season	2.8079^{***}	0.0007
Note:		*p<0.1; **p<0.05; ***p<0.01

seasonal scale (e.g., unusually hot or dry summers)—but it also captures other season-specific unobserved effects. Every year, based on the expected and factual weather, soil, and other conditions, farmers make their decisions on planting dates and volumes, crop rotation, selected hybrids, fertilizing rates, and harvesting time—among many other crop-management decisions. Many of these decisions derive from external, also season-specific, potentially adverse factors, such as hail, weeds, or disease. Growing-season clustering accounts for these factors, together with the overall characteristics of the market in a particular year and the extent to which the market can (or can not) anticipate the harvest of that season. We therefore choose to model these unobserved factors with growing-season fixed effects.

Having developed the intuition for including the selected cluster effects in one or another form, we next test our intuition formally. Because our data set is not a standard panel data set with group and time dimensions, we test for the presence of correlations for all four possible ways of clustering. Table 2.5 summarizes the Wooldridge and Hausman test statistics. In agreement with our expectation, the first test statistic indicates within-cluster correlations of some kind for all but the month-of-issue groups, suggesting that the model should account for month-of-forecast, regressor, and growing-season clustering.

The results for the second statistic are not fully consistent with this conclusion: they feature random effects over fixed effects for all but month-of-issue clustering. This test, however, relies on some strong assumptions, such as the absence of within-group correlations in the residual $\eta_{u,g}$, and therefore is not considered robust in the presence of arbitrary correlation structures (Cameron and Miller, 2015). We therefore choose to rely on the first statistic and, as discussed, implement two adjustments for the model—the fixed effects over the growing seasons and the cluster-robust standard errors with respect to the month of forecasts. The latter provide 176 clusters—a number large enough in comparison to the individual cluster size (12 observations), which makes asymptotic cluster-robust inference applicable.

Table 2.6: Results of the simple regression (2.4) for the OLS and season fixed effects (FE) estimators, with standard and cluster-robust standard errors (CRSE) with respect to the months of forecasts.

	Dependent variable:			
	Tratio			
	OLS	OLS, CRSE	FE	FE, CRSE
	(1)	(2)	(3)	(4)
R1	−0.0042*	−0.0042	−0.0040	−0.0040
	(0.0025)	(0.0053)	(0.0025)	(0.0051)
Constant	1.0046***	1.0046***		
	(0.0001)	(0.0004)		
Observations	2,009	2,009	2,009	2,009
R^2	0.0014	0.0014	0.0013	0.0013
Adjusted R^2	0.0009	0.0009	−0.0097	−0.0097
F Statistic	2.7675*	2.7675*	2.6291	2.6291
Note:				*p<0.1; **p<0.05; ***p<0.01

2.3.3 Results for the simple regression

Table 2.6 compares the OLS and the FE estimators of the simple regression (2.4), each with regular and cluster-robust standard errors. The results do not support any significant contribution of the returns to the errors of climate forecasts. The p-value of the coefficient for the regressor $R1$ computed under the OLS estimation is below 10 percent. However, the Wooldridge test statistic from Table 2.5 indicates that the OLS estimator is potentially inconsistent, and even inefficient. The adjusted standard errors, robust to clustering by month of forecast, appear more than two times larger. Introducing the fixed effects for growing seasons leaves both the value of the coefficient and its standard errors practically unchanged.

The coefficient of determination is in general expected to be low because the model relates highly uncertain returns to a hardly predictable part of the weather realization. Yet the R-squared statistic in Table 2.6 shows a diminutive amount of variance explained by the model—around 0.1 percent; the fact that the adjusted R-squared turns negative when the fixed effects are introduced further suggests the presence of too many regressors, with the amount of the explained variance close to zero. This model therefore does not pass the F-test for overall significance of the model.

Table 2.7: Results of the regression in (2.12) of the real temperature index on the forecast index.

	Dependent variable:
	Trind
Tfind	0.8529***
	(0.0041)
Constant	41.3146***
	(1.1953)
Observations	2,016
R^2	0.9550
Adjusted R^2	0.9550
Residual Std. Error	1.3896
F Statistic	42,783.5200***
Note:	*p<0.1; **p<0.05; ***p<0.01

2.3.4 Residual regression

An alternative way to formulate the question is to examine whether the returns can directly explain the weather over the growing seasons, in addition to the explanation power of forecasts. This would transform the estimation of the model in (2.4) into a two-stage estimation—first regressing the real temperature index on the forecast index,

$$Trind_m = \beta_0 + \beta_1 Tfind_m + r_m, \tag{2.12}$$

and then regressing the obtained residuals on the futures returns,

$$r_m = \gamma_0 + \boldsymbol{R'_m}\boldsymbol{\gamma} + \varepsilon_m, \tag{2.13}$$

to see if the latter contribute a significant amount of information to what was already explained by the forecast index in (2.12).

From the first step (Table 2.7) we find that the forecasts explain over 95 percent of the total variation in the observed temperatures. Table 2.8 summarizes the results for the residual regression (2.4) for the four discussed methods of estimation (the OLS and FE estimator, with the regular and cluster-robust standard errors). Clustering by the months of forecast, here too, increases the standard errors nearly twofold, while the presence of the season-specific fixed effects does not affect the results noticeably. Based on the individual t-tests of the coefficients and the overall F-test of the models, none of the four estimators suggests significant results for the model in (2.13).

Table 2.8: Results of the residual regression (2.13) for the OLS and season fixed effects (FE) estimators, with standard and cluster-robust standard errors (CRSE) with respect to the months of forecasts.

	Dependent variable:			
	residuals			
	OLS	OLS, CRSE	FE	FE, CRSE
	(1)	(2)	(3)	(4)
R1	0.7253	0.7253	0.7282	0.7282
	(0.5720)	(1.1472)	(0.5357)	(0.9679)
Constant	−0.0009	−0.0009		
	(0.0310)	(0.1005)		
Observations	2,009	2,009	2,009	2,009
R^2	0.0008	0.0008	0.0009	0.0009
Adjusted R^2	0.0003	0.0003	−0.0101	−0.0101
F Statistic	1.6078	1.6078	1.8479	1.8479
Note:				*p<0.1; **p<0.05; ***p<0.01

2.3.5 Refining the fundamental information

In this section we refine our data set following Boudoukh et al. (2007), who readdress Roll's "Orange juice paper" (Roll, 1984) and suggest that if the fundamental information is defined better it is much more strongly identified in the futures prices. In their case, focusing on a small subsample of the period crucial for orange plant growth and switching to flexible functional forms raises the R-squared from 1.5 percent to 48 percent. With regard to our data, we focus on the closest growing season and explore the predictive power of the models that differentiate the forecasts based on their leads. We find that only for the nearest lead a statistically significant relationship can be identified.

The current period of forecast

With a similar intuition—that of filtering out the information irrelevant for the financial markets—we do not expect the futures markets to perform equally well for the closer and the further seasons. Among the forecasts issued in a particular month some cover the closest (next) growing season and some might refer to the following growing season. In the group of forecasts issued in May, for example, six forecasts—for the months from May until October—will refer to that very growing season while the other two will refer to March and April of the following year. The futures trading the harvest of this latter growing season are scheduled for December the following year; preceding this contract there are five contracts trading the harvest of the current growing season and two con-

Table 2.9: Results of the simple regression (2.4) for the subset of the data covering only the current growing season. The four columns are the OLS and season fixed effects (FE) estimators, with standard and cluster-robust standard errors (CRSE) with respect to the months of forecasts.

	Dependent variable:			
	Tratio			
	OLS	OLS, CRSE	FE	FE, CRSE
	(1)	(2)	(3)	(4)
R1	−0.0037	−0.0037	−0.0048*	−0.0048
	(0.0028)	(0.0056)	(0.0028)	(0.0056)
Constant	1.0055***	1.0055***		
	(0.0002)	(0.0004)		
Observations	1,092	1,092	1,092	1,092
R^2	0.0016	0.0016	0.0028	0.0028
Adjusted R^2	0.0007	0.0007	−0.0168	−0.0168
F Statistic	1.7223	1.7223	3.0099*	3.0099*
Note:				*p<0.1; **p<0.05; ***p<0.01

tracts still trading the harvest stored from the previous year (see Figure 2.4 by way of illustration). This means that at the time of forecast issue the closest relevant contract that can be related to this forecast is only the eighth outstanding, and, given generally very low volumes of trading for the contracts after the third one (see Figure 2.7), we cannot expect the eighth contract to be efficient. We thus restrict our analysis to the forecasts for the nearest growing season and ignore any forecasts for the growing season of the following year.

As reported in Table 2.9, the selection leaves out almost half of the sample but it too does not find any evidence of significant information contributed by the returns on the futures market.

Regressions by leads for the current period

Every group of forecasts issued in a particular month has eight leads, which differ in their proximity to the month of issue. For the months that belong to growing seasons the first lead refers to the month of issue itself, and the next leads follow. For the months outside growing seasons the first lead can be several months away (up to five months away—when the forecasts are issued in November), and the other leads are correspondingly further away. With the growing season either approaching or developing, the expectations on financial markets adjust to newly received information regarding the actual growing

Table 2.10: Results of the simple regression (2.4) for the subset of the data covering only the closest growing season and only the first lead of forecast. The four columns are the OLS and season fixed effects (FE) estimators, with standard and cluster-robust standard errors (CRSE) with respect to the months of forecasts.

	Dependent variable:			
	Tratio			
	OLS	OLS, CRSE	FE	FE, CRSE
	(1)	(2)	(3)	(4)
R1	−0.0149**	−0.0149**	−0.0129**	−0.0129**
	(0.0061)	(0.0063)	(0.0060)	(0.0064)
Constant	1.0015***	1.0015***		
	(0.0004)	(0.0006)		
Observations	210	210	210	210
R^2	0.0279	0.0279	0.0240	0.0240
Adjusted R^2	0.0232	0.0232	−0.0850	−0.0850
F Statistic	5.9588**	5.9588**	4.6212**	4.6212**
Note:				*p<0.1; **p<0.05; ***p<0.01

conditions and the updated forecasts thereof. We can hence expect potentially higher predictive power of the futures markets for the closest leads, diminishing as the leads extend into the future. To test this supposition and to see if a link to returns can be identified for the periods closer to the dates of issue, we run separate regressions for the eight leads of forecasts.

Indeed, the regressions appear significant only for the first lead; we therefore omit reporting the results for further leads. As Table 2.10 reports, the coefficient for $R1$ is negative and significant at the 5 percent level. With a sample of just over 200 "observations" for the first lead, the R-squared rises to almost 3 percent in the case of OLS estimation. In the presence of the fixed effects for growing seasons the value of the coefficient for $R1$ only slightly decreases. Because almost every month serves as the first lead only once (with the exception of March, which is the first-lead forecast issued from November in the preceding year to March in the same year), clustering by the month of forecast on this restricted data set does not significantly affect the results. The sign of the coefficient for $R1$ suggests that when the forecasts on average predict too low temperatures (that is, the $Tratio$ index takes lower values), the returns are higher—a signal that the market interprets the expected warmer weather conditions as adverse and reacts with increases in price.

2.4 Discussion and further exploration of the data

The results in Section 2.3 in general do not support the main intuition developed in the first part of the paper: the regression analysis, with an exception of the nearest-lead regression, does not identify the presence of information additional to the available climate forecasts. The chosen setting is rather conservative, in that it uses the best available forecasts (in fact, hindcasts) produced by a multi-model ensemble—a system that outperforms any individual climate model. In fact, the hindcasts available to the public are in many cases already corrected for known biases exhibited by the models in their earlier runs (Kirtman et al., 2014). Our additional examination of the forecasts of *individual* models, however, leads to a similar conclusion: even though our proxy variable for the error in climate forecasts, $Tratio$, takes more extreme values when the data are not averaged across the models, there is no clear signal related to the financial market.[16]

Another feature that potentially hinders information discovery is the *frequency* of the data collected from the futures market. The advantage of using monthly data is that they align with the frequency of forecast issue and are affected to a much smaller extent by day-to-day shocks such as news releases. At the same time, for an efficient market, the effect of changes in the expected growing season conditions might be also hard to detect on scales longer than a very few days. *Market efficiency* itself is one of the conditions that our hypothesis relies on; yet, as mentioned in the earlier discussion, it is claimed but not clearly established in the literature.

Turning to agricultural factors, on which our results heavily depend, we might have to account for the timing and conditions of corn plants' growth more thoroughly. In particular, there is strong evidence that only a certain interval of high temperatures negatively affects the plant growth. To test this supposition, we additionally analyzed the areas and months where and when the surface temperature exceeded the *critical threshold* of 30℃ (303.15 K) suggested by Schlenker and Roberts (2006). However, based on the monthly data, there is no clear evidence that the markets anticipate these adverse extremes. Even though the temperatures above and below this threshold should be perceived as having antithetical effects, the results for the two separate parts of the data do not differ notably.

One supposition is that *irrigation* might alleviate the impact of extremely high temperatures. Sufficient water supply, however, is critical at the early stage of corn plants' growth, whereas temperature extremes are likely to appear much later—during ears' development stage—when additional irrigation does not necessarily compensate for reduced

[16]The results of the additional analyses mentioned in this section are in line with those obtained in the previous section. In either case, no significant connection is identified between the market returns and the errors in the climate forecasts. We therefore chose to omit all but a few obtained results of these additional regressions.

photosynthesis. In fact, less than 15 percent of corn area in the US was irrigated in the sample period.[17] We therefore do not find a significant difference between the analyses of all and non-irrigated-only areas.

Our definition of a growing season is also rather conservative. As described in Section 2.2.5, according to the biology of corn plants, not all months are equally important for their growth—July and August appear to be the critical months for the ear's formation and growth and, consequently, for the harvest of the season. At the same time, given the differences in planting and harvesting times across states and counties, the inclusion of March and October might bring too many irrelevant observations into the data set and obscure the signal from the other months. Focusing the analysis on the most *important* months, however, leaves the results surprisingly unchanged while reducing the number of observations to only few dozens.

One assumption our analysis relies on is that changes in the expected weather conditions directly transfer to signals on the futures market. To have this effect, the information has to be significant and concern the participants of this market. The changes in the weather conditions, first of all, concern the producers of the commodity; and, if these changes transform into supply shocks, they affect the market. If, however, there is some adjustment from the farmers' side—if they can adapt their crop-management decisions or simply smooth supply using existing stocks—the market might remain unaffected. In fact, the futures market might be more sensitive to stored volumes of grain—that is, to the sufficiency of stored corn to cover potential reductions in harvest. Farmers' decisions on how much to keep and how much to sell, in turn, depend on multiple factors—the yields of a given season, market trends and expectations, and, finally, available storage space and conditions. The allocation of the harvest between the new-crop December futures contract and later contracts thereby reflects the general storage strategy for that year. On average, the harvest is typically sold out by the end of April, except the share kept for own use such as livestock feed (Maier, 2004); it is therefore untypical for farmers to carry their stocks over to the next year.

The data on corn storage are available in the same database, QuickStats of USDA. Every year the values are reported as of the 1st of March, June, September, and December. Figure 2.8 places the reporting quarters on the time line and displays the dynamics of corn storage over the sample period. The reported volumes of stored corn are rather high even in the last reporting month (September) because both on-farm and off-farm stocks are counted; that is, even though almost nothing is carried over to the next growing season on farms, elevators and other buyers can adjust their own stocks.

[17] USDA Census of Agriculture, `https://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Census_of_Aquaculture/index.php`

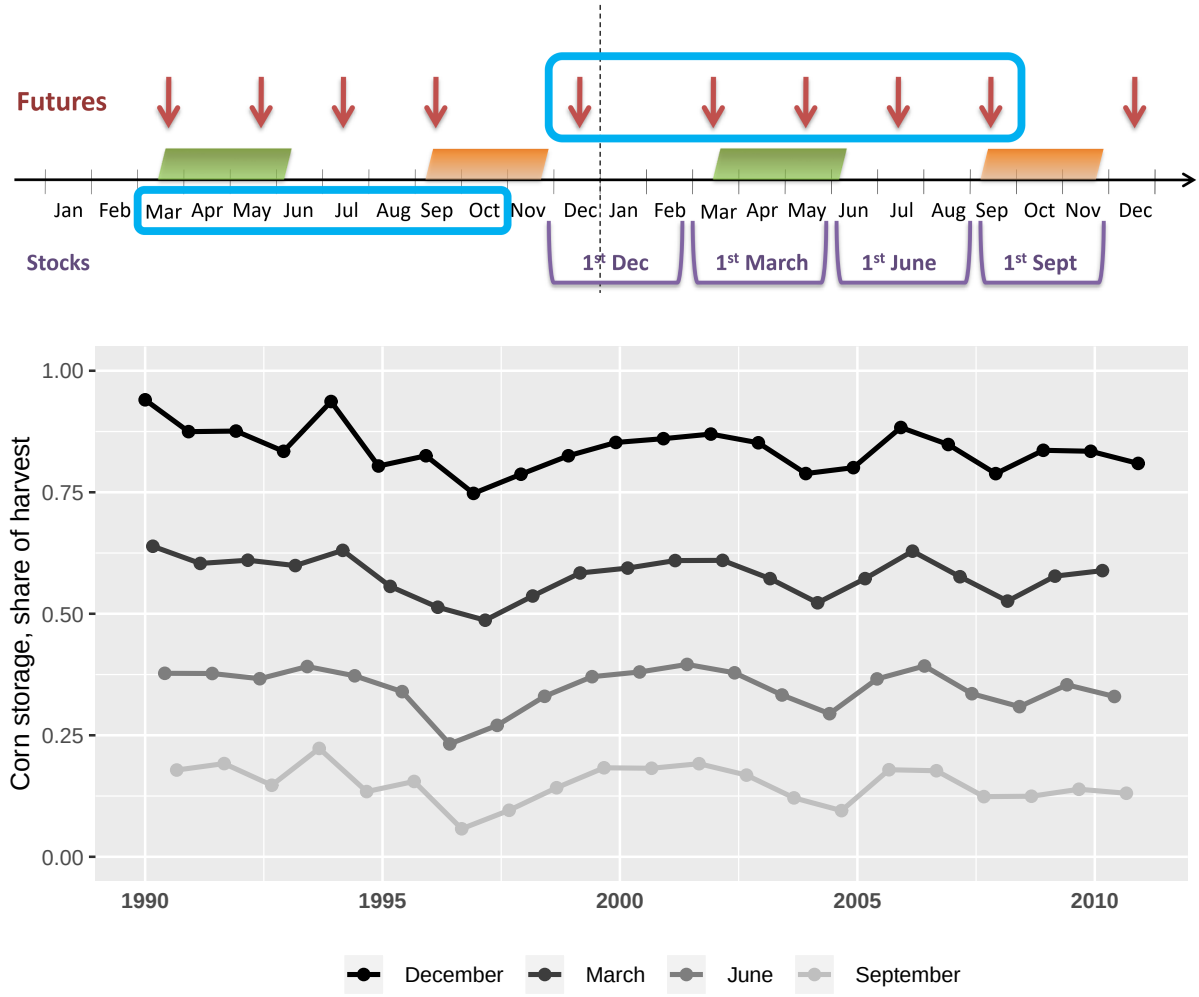


Figure 2.8: Time line relating the corn growing season, futures contracts, and report dates for remaining stocks of corn (top, analogous to Figure 2.4) and the dynamics of the stocks of corn over the sample period (bottom). The reporting periods for stocks are marked with purple "bundles" on the top panel and with different shades of gray on the bottom panel.

The results of our analysis are summarized in Table 2.11. The so-called *ex ante* stocks—the stocks reported at the time of forecasts issue—do not exhibit any adjustment with respect to forecasted weather, let alone to its future realization. These stocks, however, are highly correlated with contemporaneous returns on futures. The *ex post* stocks—the stocks at the time of weather *realization*—in turn, are strongly affected by the realization of weather conditions over the growing season; in fact, the unexpected component (weather shocks) explain 11.5% of fluctuation in these stocks. These fluctuations are not anticipated by the futures market. These results suggest that both expected and unexpected components of weather during growing seasons strongly influence corn supply; this influence, however, is not anticipated but rather absorbed by the futures market.[18]

[18]The coefficients in the regressions (5)–(7) in Table 2.11 suggest an inverse relationship between stocks and contemporaneous temperatures—warmer conditions on average are considered favorable and

Table 2.11: Results for the regressions of the shares of stored corn reported at the time of forecast issue (*Ex ante* stocks) and at the time of weather realization (*Ex post* stocks). Here $Tratio$ and $Trind$ are *ex post* regressors; $Tfind$ and $R1$ are *ex ante* regressors.

	Dependent variable:							
	Ex ante stocks, %				Ex post stocks, %			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tratio	0.3191				−9.4623***			
	(0.9446)				(0.5906)			
Trind		−0.0016				−0.0104***		
		(0.0014)				(0.0019)		
Tfind			0.0001				−0.0095***	
			(0.0008)				(0.0005)	
R1				0.7168***				0.0104
				(0.1613)				(0.1122)
Const.	0.1580	0.9289**	0.4426**	0.4644***	9.8871***	3.3720***	3.1482***	0.3596***
	(0.9489)	(0.3918)	(0.2245)	(0.0090)	(0.5932)	(0.5475)	(0.1370)	(0.0061)
N obs.	2,016	806	2,016	801	1,972	168	1,972	838
R^2	0.0001	0.0018	0.00001	0.0241	0.1153	0.1522	0.1716	0.00001
Adj.R^2	−0.0004	0.0005	−0.0005	0.0229	0.1148	0.1470	0.1711	−0.0012
RSE	0.2610	0.2555	0.2610	0.2531	0.1623	0.1604	0.1571	0.1777
F Stat.	0.1141	1.4119	0.0257	19.7397***	256.7248***	29.7894***	407.9861***	0.0085
Note:								*p<0.1; **p<0.05; ***p<0.01

Finally, even though we consider the largest corn producing economy, global market prices might have the effect of averaging out the signals from local markets and take into account potential substitution by other countries.

2.5 Conclusion

The well-established evidence of ongoing climate change suggests strong adverse effects on vulnerable economic sectors, including agricultural food production. The general warming trend covers main crop growing areas, with temperatures over the growing season turning harmful more frequently in recent decades. Changing weather patterns subsequently affect crop growing practices and might induce supply shocks on the markets for agricultural commodities. We therefore suggest that it is strongly in the interests of par-

the stocks are lower in anticipation of richer harvest. An analysis of "too hot" temperatures (above 30℃, Schlenker and Roberts 2006) alone, however, reveals the expected reverse relationship in which unanticipated adverse heat induces higher stocks levels.

ticipants in the financial markets for agricultural commodities to obtain the best possible information about future climate conditions—information that might help improve their strategies on the market. We therefore expect efficient financial markets to capture in their prices all information available to their participants with regard to the expected weather conditions in the coming growing seasons.

This paper attempts to find evidence of the proposed relationship. Using the example of corn as a crucial crop we study whether the futures markets for this commodity reflect any information about future climate conditions on top of the best publicly available forecasts. In particular, we examine the relation between the returns on corn futures and the systematic errors in the climate forecasts of a multi-model ensemble over the areas in which a major part of the crop is grown.

Despite the evidence of the efficiency of the corn futures markets in the literature, no clear signal for weather expectations more accurate than those of the general public can be identified in our data set. Filtering the parts of the data most relevant for the participants of futures markets, à la Boudoukh et al. (2007), as well as additional more detailed analyses do not reveal any significant effect beyond the nearest forecast. We conclude that the proposed framework does not accurately describe the mechanism of interaction between climate conditions, agricultural production, and the trading of the produced commodities on financial markets.

Essay 3

Global Sensitivity Analysis for Integrated Assessment Modeling

Global Sensitivity Analysis for Integrated Assessment Modeling[1]

Alena Miftakhova, University of Zurich

Abstract

Integrated assessment modeling studies the nexus between the systems of the climate and the economy, both known for their high complexity and vast uncertainty. One key question is how sensitive climate policy inference produced by such models is to the uncertainty in their initial assumptions and the results of calibrations. Despite the broad literature on the topic—rich in both single-model and multi-model sensitivity analyses—universal, well-established practices for analyzing the uncertainty in models' outcome are still missing. In this paper we argue for structured global sensitivity analysis (GSA) as an indispensable routine in climate–economic modeling. We apply a high-efficiency GSA method based on polynomial chaos expansions to the most commonly employed integrated assessment model (IAM), DICE. Our analysis provides two key insights. First, the subjective preselection of a subset of parameters of interest might omit the most influential ones. Second, the opposite strategy—one of pooling all parameters together and considering the model as a "black box"—might distort sensitivity indices. The best practice is thus to consider all exogenous parameters but adjust their set such that the fundamental relations within the model's structure are respected. The methodology of efficient GSA provides a clear, comprehensive decomposition of the uncertainty in a model's output while minimizing computational costs, and hence is easily applicable to IAMs of higher complexity.

[1]We are indebted to Daniel Harenberg, Karl Schmedders, and Reyer Gerlach for insightful discussions on the subject. We thank the participants at the EAERE Winter School 2018, EAERE-FEEM-VIU European Summer School 2018, and the 11th Annual Meeting of the IAMC (2018) for helpful comments. We are grateful to Dave Brooks for excellent editorial support. Alena Miftakhova gratefully acknowledges financial support from the Swiss National Science Foundation.

3.1 Introduction

Integrated assessment models (IAMs) are meant to provide scientific support to policy makers in dealing with the complex, global phenomenon of climate change. Aimed at pinning optimal climate policy down, these models have gained the trust of the authorities for the practical guidance they provide regarding timely climate actions (EPA, 2014, 2016). The inference implied by the solutions of IAMs often directly transfers into climate policy advice; yet the factors that drive these solutions remain obscure as most of these models are too complex to track down the source of their inference. Despite a clear call for a better communication of the form, magnitude, and sources of the uncertainty associated with the reported "best estimates" for climate actions (Webster, 2003; Diaz and Moore, 2017), thorough, inclusive practices of sensitivity analysis are uncommon in integrated assessment studies. In this paper we address the task of establishing a well-founded practice of comprehensive sensitivity analysis in climate–economic modeling. We take a probabilistic approach and employ the methodology of global sensitivity analysis (GSA) to identify the main sources of uncertainty in an IAM and their relative strength of influence on the model's outcome. The practice of global sensitivity analysis, on the one hand, verifies the robustness of models' inference and, on the other hand, informs the modeler about their applicability to ever-changing, uncertain, dynamic systems such as the economy or the climate (Hawkins et al., 2016; Atkinson et al., 2009). An application to a basic IAM is enough to demonstrate the importance of inclusive sensitivity analysis as opposed to widespread selective practices.

The forces that drive the inference in IAMs and cause the disagreement on the optimal climate policy have been highly debated in the literature. One key value that serves as a reference point for climate policies produced by IAMs is the estimated *social cost of carbon* (SCC). The SCC refers to the present value of all future damages from a marginal (one ton) increase in carbon dioxide emissions. The range of the estimated values for the SCC in the literature is enormous and heavily depends on the models' assumptions (Arent et al., 2014). This phenomenon gave rise to the literature that tries do discover the rules for optimal climate policy in IAMs by the means of analytical derivations (Golosov et al., 2014; Rezai and Van der Ploeg, 2016; van den Bijgaart et al., 2016; Dietz and Venmans, 2017; van der Ploeg and Rezai, 2019). The studies identify the key assumptions that determine the value for the SCC in the integrated assessment framework and group them into the *economic* factors—the discount rate, the intergenerational inequality aversion, and the rates of growth for the economy and population; the *climate* factors—the short-term rate of depreciation of CO_2 in the atmosphere and the rate of temperature adjustment to its long-term equilibrium; and the *interaction* factor—the severity of damage to economy from rising temperature.

These theoretical results are invaluable for understanding the general IAM framework. In practice, however, many IAMs are much more detailed and therefore too sophisticated to enjoy analytical solutions. Here, we aim to demonstrate the application of comprehensive yet affordable sensitivity analysis that reveals the factors that contribute to the variation in a model's results. We intentionally apply the method to a simple IAM, so that the results are comparable with the analytical inference.

Point estimates, which neglect uncertainty, are long since considered dissatisfactory in economic modeling (Leamer, 1985; Canova, 1994, 1995; Pindyck, 2013). IAMs, which connect the economy with the physical climate system, face the even greater challenge of accounting for yet more uncertainties while staying in agreement with scientific knowledge. Typical trade-offs faced by an IAM modeler—the formulation of models, the simplification of the structures of modeled systems, the choice of the components to include or to leave out, and the parameterization of the modeled processes—have a tremendous effect on the consequent inference (Calel and Stainforth, 2017).

In an attempt to reach more robust conclusions, almost all recent IAM studies involve an analysis of uncertainty in one form or another. Among numerous methods in the literature, *local methods*, which focus on particular points in parameter space, have gained popularity thanks to their simplicity (Saltelli and Annoni, 2010). These methods by definition suffer from dependence on the parameters' values chosen for examination, and therefore might overlook important regions of interest. Even when the analysis of an IAM covers the entire range of possible parameter values, the set of parameters in question is often limited to only a very small number that "intuitively" matter according to expert judgment or the results of preceding work (Kann and Weynant, 2000; Peterson, 2006). In multi-model studies, considering a *subset* of common parameters of interest is a necessary restriction, whereas for single-model analysis such practice is yet another source of bias, as expert or experience-based judgment might turn out to be inconsistent (Millner et al., 2013; Wesselink et al., 2015; Saltelli et al., 2010). Last but not least, the interaction effects of parameters in IAMs and their nonlinear impact—often undetectable using local methods—are not to be neglected (Butler et al., 2014; Anderson et al., 2014).

Global methods, in contrast, offer a comprehensive picture of uncertainty in a model's inference by, first, exploring the entire parameter space and, second, accounting for interactions and nonlinearities in parameters' effects on the output. Many of these methods are *variance-based*—they decompose the variance of output into the shares attributed to each input parameter and their combinations.[2] The analysis thereby ranks all input parameters according to their relative importance, individually and in their interplay. As such, this ranking covers the entire specified uncertainty domain—an advantage that

[2]The reader is referred to Borgonovo and Plischke (2016) for a frontline overview of both local and global sensitivity analysis methods and to Ghanem et al. (2017) for a practical guidance thereto.

comes at a high computational cost of thousands of model runs (Saltelli et al., 2008).

In this paper we focus on reducing these costs by the means of a specific GSA method that produces a full variance-based representation of uncertainty in a model much more efficiently than alternative techniques do. The method involves *polynomial chaos expansions* (PCEs), which approximate the outcome of a model with a sum of multivariate orthogonal polynomials. The use of PCE for sensitivity analysis was introduced by Sudret (2008) and has been frequently applied in the engineering literature; Harenberg et al. (2019) bring the method to economics. The most important property of the approximation is that it produces a full variance-based representation of uncertainty in the model from a relatively small number of model runs. We therefore advocate the use of the method in integrated assessment modeling, where parameter uncertainty is profound but models' complexity often limits the affordable number of model evaluations.

We demonstrate the application of the method on the most widely employed IAM, DICE (Nordhaus, 2008). When applied to a selection of parameters, the method produces results consistent with the conclusions of Nordhaus (2008). Full analysis, however, changes the ranking dramatically, with the most influential parameters being among those omitted from the aforementioned restricted setting. We also show that caution is needed when analyzing the full set of parameters because the credibility of the results relies on the *independence* assumption. In the case of DICE this implies reformulating the GSA setting such that it sustains the fundamental relationships embedded into the model's structure. Our work therefore promotes careful, informed inclusion of all potentially important parameters into GSA.

To date, only few IAM studies consider global sensitivity analysis viable. Anderson et al. (2014) use transformation-invariant global sensitivity indices to prioritize all exogenous parameters of DICE. Butler et al. (2014) via variance decomposition analyze the vulnerability of policy scenarios run in DICE to the uncertainty in a selected set of its initial parameters. Both studies follow the common practice of generating large Monte Carlo samples of input–output data to estimate sensitivity indices. The key advantage of the method applied in this work—the use of polynomial chaos expansions for both constructing a meta-model and deriving the importance indices—makes it much more efficient and therefore affordable for more complex IAMs.

The rest of the paper is structured as follows. Section 3.2 outlines the structure of the DICE model, identifies the major factors driving its results by the means of local sensitivity analysis, and presents the analytical results in the literature. Section 3.3 explains, step by step, the methodology of the efficient method of global sensitivity analysis that we use and illustrates its advantages. Section 3.4 presents the results of the PCE-based GSA applied to DICE, comparing restricted and complete settings and listing the relationships

to be reformulated. Section 3.5 concludes.

3.2 DICE model structure

This section begins with the outline of the key components of the DICE model; for a detailed description the reader is referred to Nordhaus (2008). We next illustrate the traditional, local approach to the sensitivity analysis of an integrated assessment model, and DICE in particular, in a selected- and full-set settings. The analysis is accompanied by the recent analytical results in the literature.

3.2.1 Economy and climate in DICE

DICE is a neoclassical growth model that resolves a trade-off between consumption, investment, and emissions reduction. Given a discount rate ρ, the model maximizes total social welfare over its time span,

$$W = \sum_{t=1}^{Tmax} (1+\rho)^{-t} U(C_t, L_t), \tag{3.1}$$

with a constant-elasticity utility function of consumption, C_t, and population, L_t,

$$U_t = \frac{(C_t/L_t)^{(1-\alpha)} - 1}{1-\alpha} L_t. \tag{3.2}$$

The production function takes the standard Cobb–Douglas form with Hicks-neutral total factor productivity (TFP) A_t, labor L_t, and capital K_t as input factors. Total output Y_t is reduced by the cost of mitigation—a power function of emission reduction policy μ—and by the damage induced by the climate change, Ω_t,

$$Y_t = \Omega_t (1 - \theta_1 \mu_t^{\theta_2}) A_t L_t^{1-\gamma} K_t^{\gamma}. \tag{3.3}$$

The first term of equation (3.3)—the reduction in economic output caused by climate change—is a function of the atmospheric temperature anomaly T_A (measured as the increase in the atmospheric temperature from the preindustrial level),

$$\Omega_t = \frac{1}{1 + a_1 T_{A_t} + a_2 T_{A_t}^{a_3}}. \tag{3.4}$$

By default the damage function is assumed to be of a reduced quadratic form with $a_3 = 2$ and $a_1 = 0$.

The reduced, net output is distributed between consumption and investment I_t, which directly transforms into the capital of the following period. Given a depreciation rate δ_K,

capital accumulates according to

$$K_t = (1 - \delta_K)K_{t-1} + I_{t-1}, \tag{3.5a}$$

$$I_t = Y_t - C_t, \tag{3.5b}$$

while the growth paths for labor and TFP are given exogenously.

Every period, the production of goods affects the climate system by inducing anthropogenic CO_2 emissions, at a rate determined by the emissions intensity σ_t and reduced by the chosen rate of mitigation μ_t,

$$E_{ind_t} = \sigma_t(1 - \mu_t)A_t L_t^{1-\gamma} K_t^{\gamma}. \tag{3.6}$$

Together with natural emissions E_{tree}, anthropogenic emissions comprise the total CO_2 emissions level E_t,

$$E_t = E_{ind_t} + E_{tree_t}. \tag{3.7}$$

The core of the *climate* module of the model is its carbon cycle, which is represented by a three-layer model of atmospheric, upper layer, and lower ocean CO_2 concentrations (M_{AT_t}, M_{U_t}, and M_{L_t}, correspondingly), with transitions among these layers defined by the following system of equations:

$$M_{AT_{t+1}} = E_t + b_{11}M_{AT_t} + b_{21}M_{U_t}, \tag{3.8a}$$

$$M_{U_{t+1}} = b_{12}M_{AT_t} + b_{22}M_{U_t} + b_{32}M_{L_t}, \tag{3.8b}$$

$$M_{L_{t+1}} = b_{23}M_{U_t} + b_{33}M_{L_t}. \tag{3.8c}$$

According to this system, CO_2 emissions first enter the atmosphere and in the course of following periods get partially absorbed by the upper layer; from there they are slowly transmitted to the lower ocean—until the system reaches new equilibrium proportions of concentrations among the three layers.

The net change in atmospheric CO_2 concentrations indices the net energy flow to the atmosphere (called radiative forcing, F_t),

$$F_t = \eta \log_2 \left(\frac{M_{AT_t} + M_{AT_{t+1}}}{2M_{AT_{1750}}} \right) + F_{EX_t}, \tag{3.9}$$

where F_{EX_t} is exogenous radiative forcing. This energy causes a change in the atmospheric temperature, T_{A_t}, and, subsequently, the ocean temperature, T_{O_t},

$$T_{A_{t+1}} = T_{A_t} + C_1 \left(F_{t+1} - \frac{\eta}{\lambda} T_{A_t} - C_3 \left(T_{A_t} - T_{O_t} \right) \right), \tag{3.10a}$$

$$T_{O_{t+1}} = T_{O_t} + C_4 \left(T_{A_t} - T_{O_t} \right). \tag{3.10b}$$

Table 3.1: Parameters selected for the analysis and their distributions (from Nordhaus, 2008). The column *Mean* corresponds to the mean of the normal distribution; the column *Std* to its standard deviation.

Parameter	**Definition**	**Mean**	**Std**
a_2	Damage function coefficient	0.0028	0.0013
b_{12}	Carbon cycle transition	0.189	0.017
g_{A0}	Initial growth rate of TFP	0.092	0.040
$g_{\sigma 0}$	Initial change of decarbonization	-0.07	0.02
$CCum$	Maximum extraction of fossil fuels	6,000	1,200
p_{BACK}	Cost of backstop technology	1.17	0.468
L_{ASYM}	Asymptotic population	8,600	1,892
λ	Equilibrium climate sensitivity	3.00	1.11

As specified above, the change in the atmospheric temperature has a subsequent negative feedback on the output of the economy.

For the sake of comparability with the existing literature, we use the version *DICE delta version 8*.[3] This version of DICE covers a time span of 60 periods, with an increment of 10 years, and operates on the aggregated global level; the model has 51 parameters listed in Table 3.A. Additionally, Section 3.A of the Appendix lists the 15 parameters that are never considered in this study due to their irrelevance for the outcome of the model.

3.2.2 Local sensitivity analyses of DICE

The first sensitivity analyses of DICE are local one-at-a-time analyses of selected parameters, accompanied by Monte Carlo simulations from the distributions assumed for these parameters (Nordhaus, 1993, 2008). These analyses essentially measure the magnitude of the changes in model's major output quantities—the social cost of carbon (SCC), total CO_2 emissions, and atmospheric temperature anomaly—induced by individual deviations in parameters' values. The simulations, in turn, estimate the distribution of these outcome quantities when the uncertainties in the selected parameters are considered (sampled) jointly.

The experiment in Nordhaus (2008) involves eight parameters that presumably have the strongest effect on the outcome of the model; their assumed normal distributions are summarized in Table 3.1. In this analysis of the eight selected parameters, the SCC in 2005 is affected the most by the linear damage coefficient, a_2, climate sensitivity, λ, and the initial growth rate of technology, g_{A_0}. The asymptotic size of population, L_{ASYM} has a smaller effect; the impact of the remaining four parameters is insignificant.

[3]This version of DICE is retrieved from `http://www.econ.yale.edu/~nordhaus/homepage/homepage/DICE_delta_v8_YUP_book_short_noexclude.GMS` last accessed on 21 November 2018.

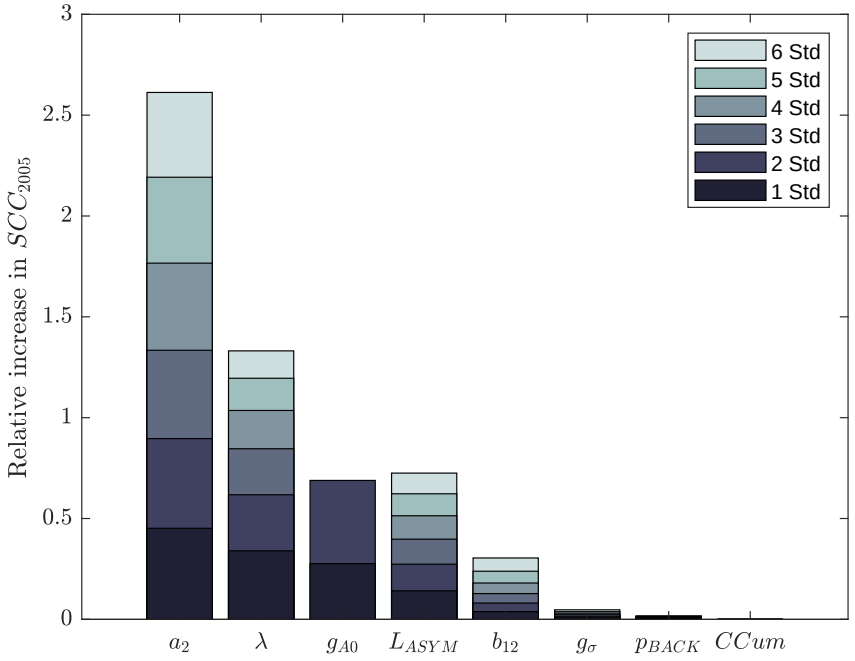


Figure 3.1: One-at-a-time effects in the setting analogous to Nordhaus (2008)—changes in the social cost of carbon in 2005 induced by changes in parameters' values to 1–6 standard deviations from their means. The parameters are sorted by the magnitude of the change induced by one-standard-deviation step. In each case, the effects of changes in the direction that *increases* SCC are reported.

Using the same specification, we calculate the relative changes in model's outcome (the social cost of carbon in 2005) from varying each parameter by 1 to 6 standard deviations from its mean. We perform this and all other experiments in the *optimal*, not the baseline, setting of DICE because our interest lies in the effect on the optimal climate policy, as opposed to business as usual scenario. Figure 3.1 shows the results for the SCC.[4]

As a more general experiment, we consider the full set of the parameters in DICE (with trivial exceptions listed in Section 3.A) and vary their values one-at-a-time. To exclude any subjective probabilistic assumptions, we vary each parameter Θ_i on the interval $\left[0.5 \times \bar{\Theta}_i, 1.5 \times \bar{\Theta}_i\right]$ with the step $0.1 \times \bar{\Theta}_i$, where $\bar{\Theta}_i$ is the default value of the parameter Θ_i in DICE.

Figure 3.2 displays the results of this general OAT analysis[5] and suggests three key insights that *undermine* the setting above. Most noticeably, three parameters with the most significant impact on the SCC—the exponent of the damage function, a_3, the capital elasticity, γ, and the elasticity of marginal utility of consumption, α—are not included

[4]With its default optimization settings the program does not solve the model for the values of g_{A_0} for more than two standard deviations to the left of its mean. The value of g_{A_0} in this case turns negative, which, according to Nordhaus (2008), should be excluded from the analysis. We therefore omit reporting the results for the cases of g_{A_0} decreased by 3 to 6 standard deviations from its mean.

[5]For some of the solutions to be feasible we had to relax the constraint on the initial value of capital, $K_0 \geq 100$.

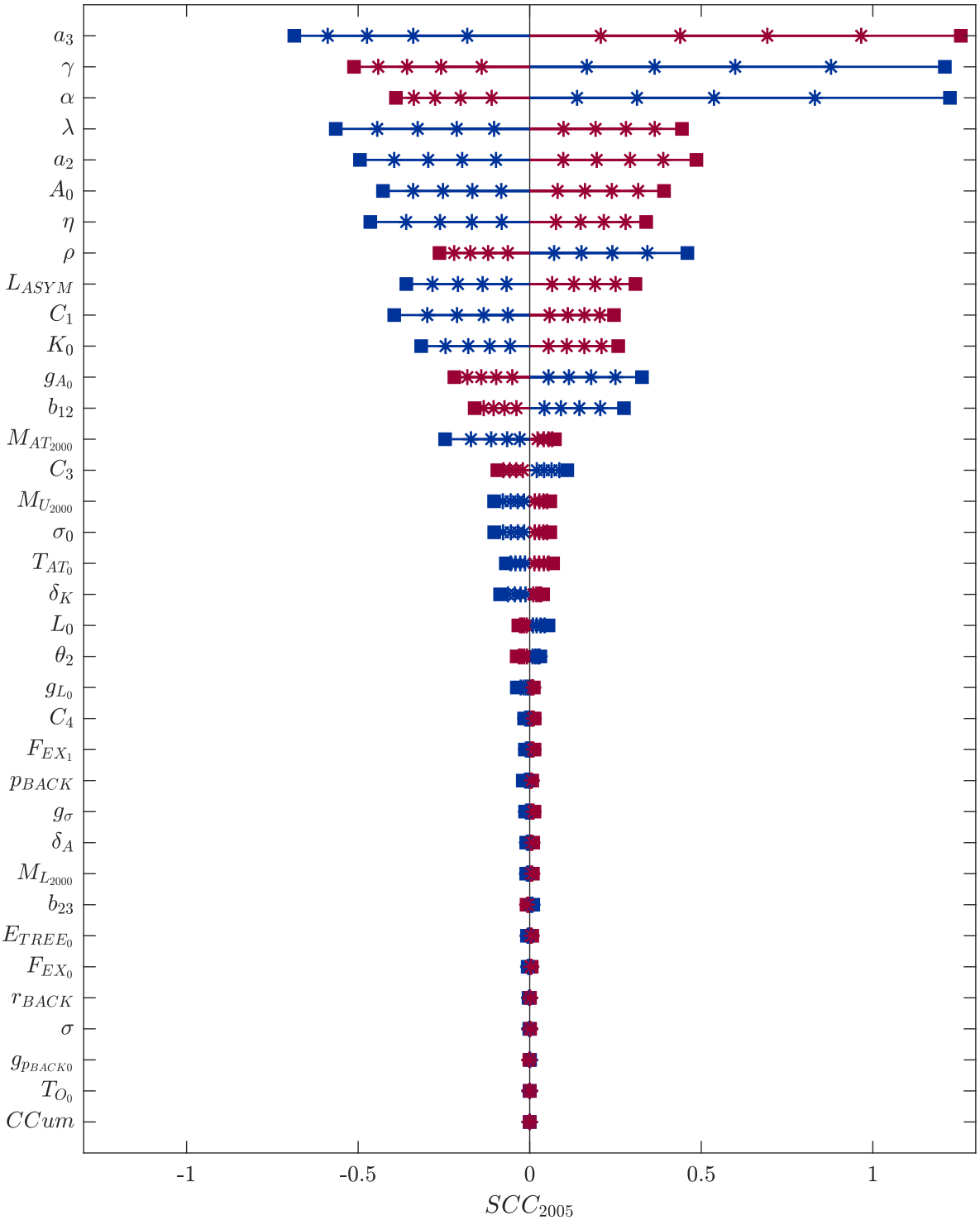


Figure 3.2: One-at-a-time effects of varying DICE parameters, sorted by the total magnitude of their positive and negative effects. The stars sequentially mark the effects of 10%, 20%, 30%, and 40% changes in parameters' values; the squares the effect of 50% changes. The effects of the increases in parameters' values are shown with red bars; the effects of decreases with blue bars.

in the analysis above. Second, many more than 8 parameters potentially have an impact on SCC. Apart from the highly debated discount rate, ρ, and initial conditions for capital, K_0, and technology, A_0, the equilibrium radiative forcing from a doubling of CO_2 concentrations, η, and the rate of long-term temperature adjustment to its equilibrium level, C_1, are the obvious candidates for sensitivity analysis. Finally, the effects of the parameters are *not* linear: for many parameters (including g_{A_0} and b_{12} selected in Nordhaus (2008) as influential) the magnitudes of positive and negative changes in the SCC from varying these parameters in opposite directions differ significantly.

3.2.3 Theoretical results in the literature

Sensitivity analysis is one common way to explore the factors that contribute to the variation in the output of a model. For complex nonlinear models—which mostly do not allow for analytical solutions—it is the only feasible way to identify the main drivers behind models' results. Lately, much more *analytical* work has been done in the integrated assessment literature to explore the factors that drive the social cost of carbon in general climate–economic frameworks. In their pioneering work, Golosov et al. (2014) derive a closed-form formula for the social cost of carbon in a dynamic stochastic general equilibrium (DSGE) model that features damages from climate change. They suggest that, under certain assumptions—such as logarithmic utility and a constant savings rate—the optimal carbon tax is *proportional* to the output of the economy. In their general setting, the three major *determinants* of the tax's share are the discount rate, the elasticity of damage function with respect to output, and the depreciation rate of CO_2 in the atmosphere.

Van den Bijgaart et al. (2016) extend the analysis to a more general structure—an economy that follows a balanced-growth path but is not restricted to *unitary* elasticities of marginal utility and of damages with respect to output. The model additionally allows for non-constant elasticity of damages with respect to temperature. According to their analysis, the key *economic* parameters driving the SCC are the discount rate, ρ, elasticity of marginal utility, η, and the growth rate of the population, l, and of consumption, g. The parameters of the *climate* system that affect the SCC are climate sensitivity, c, the depreciation rate of CO_2 in the atmosphere, δ_S, and the rate of adjustment to the long-run equilibrium temperature level, ε. The rest of the relevant factors are those that determine the shape of the damage function: the reference damage to the economy at 1 degree of temperature increase, ω, and the elasticity of damage function with respect to temperature, Ψ, and to output, ξ. The authors conclude that the social cost of carbon in a generally specified economy with gross output Y and population L can be well

approximated by the formula

$$SCC_t = \frac{1.3\omega c^{\Psi}}{m} \frac{1}{\delta_S + \sigma} \frac{\varepsilon}{\varepsilon + \sigma} \left(\frac{Y_t}{L_t \bar{y}} \right)^{\xi - 1} Y_t, \tag{3.11a}$$

$$\sigma = \rho + (\eta - \xi)g - l, \tag{3.11b}$$

where m is the preindustrial level of CO_2 concentrations, and $\bar{y}$ is reference income per capita corresponding to the temperature rise of 1 degree and the damage ω. The authors demonstrate that, when applied to the DICE model, the formula captures the main drivers of the SCC and thereby provides the analytical base for the discussion on the factors that define optimal carbon tax in this and other IAMs.

Rezai and Van der Ploeg (2016) and van der Ploeg and Rezai (2019) modify the model to include additional features (e.g., hyperbolic discounting, temperature lag, and climate damages not proportional to GDP) and confirm the proportionality of the optimal carbon tax to the output; they also identify the importance of the discount rate, the growth rates of economy and population, the short-term dynamics of the CO_2 concentration, the rate of long-term adjustment of the atmospheric temperature to its equilibrium rate, and the relative damage to economy. In the model of Dietz and Venmans (2017) the same economic factors drive the social cost of carbon; in regard to climate, however, the authors point out the fallacy of most IAMs to account for the saturation of carbon sink and, after proper modifications, name transient climate response to cumulative carbon emissions as a key climate factor.

Notably, the studies tend to use simple (*local*) sensitivity analysis methods. Dietz and Venmans (2017) use OAT to map the sensitivity of optimal peak warming to various values of the parameters the define it. Rezai and Van der Ploeg (2016) use scenario analysis with alternative assumption on climate damages, population growth, and economic growth.

Yet the model that stays the closest to the original DICE is that of van den Bijgaart et al. (2016). Because DICE assumes damages from climate change proportional to income (that is, with *unit* elasticity ξ), the formula above reduces to

$$SCC_t = \frac{1.3\omega c^{\Psi}}{m} \frac{1}{\delta_S + \sigma} \frac{\varepsilon}{\varepsilon + \sigma} Y_t, \tag{3.12a}$$

$$\sigma = \rho + (\eta - 1)g - l. \tag{3.12b}$$

The parameters involved in the formula are directly present in DICE, with only three exceptions. The growth rate of the population, l, is determined by its initial and asymptotic values and the initial growth; the growth rate of consumption, g, is an endogenous

Table 3.2: Correspondence of the parameters included in the SCC formulas (3.12) and those in DICE.

Parameter	**Definition**	**DICE analog**
ω	relative damage at the temperature rise of 1℃	$\frac{a_2}{1+a_2}$
c	climate sensitivity	λ
Ψ	damage–temperature elasticity	a_3
δ_S	rate of depreciation of atmospheric CO_2	b_{12}
ε	rate of adjustment to long-run equilibrium temperature	C_1
ρ	time discount rate	ρ
η	elasticity of marginal utility of consumption	α
l	population growth rate	L_0,L_{ASYM}, g_{L_0}
g	consumption growth rate	endogenous
m	preindustrial concentrations of CO_2	set to 596.4

variable that accompanies the model's solution; the preindustrial level of CO_2 concentrations is embedded in the code as a constant. Table 3.2 lists the parameters in the formula with their analogs in DICE.

Based on these analytical results, in our analysis we can expect the highest significance indices for the parameters that participate in the formulas (3.12). Deviation from the formula would be a sign of either our more detailed analysis revealing that some important factors are omitted from (3.12) or spurious significance of some parameters that is in fact driven by one of those included in the formula—the potential problem addressed in Section 3.4.4.

3.3 PCE-based global sensitivity analysis

GSA methods take a probabilistic approach in that they view any model as a function of the random vector of its input parameters, with probability distributions assigned to them prior to the analysis. Once the distributions are specified, the uncertainty of the input parameters can be transferred to the output by evaluating the model on a sample of points drawn from this parameter space. The resulting sample helps visualize the (otherwise unknown) distribution of the output and estimate its moments. As a next step, variance-based GSA conveniently apportions the total variance in the output between individual parameters and their interactions. The strength of a parameter's influence, also called its sensitivity index, is hence defined as its share of the total variation of the output. The rest of this section presents the methodology in more detail, including the use of polynomial chaos expansions in the estimation of the sensitivity indices for a model's parameters.

3.3.1 Uncertainty propagation

From the perspective of GSA, the model of interest is a *mapping* from a random vector of its exogenous parameters $\mathbf{\Theta}$ onto an endogenous output Y,

$$Y = \mathcal{M}(\mathbf{\Theta}), \tag{3.13}$$

where $\mathcal{M} : \mathbf{\Theta} \mapsto Y$, with $\mathbf{\Theta} \in \mathbb{R}^M$ and $Y \in \mathbb{R}$. Each parameter in $\mathbf{\Theta}$ is characterized by a *probability density function* (PDF) $f_{\Theta_i}, i = 1, \cdots, M$. The assumptions of appropriate probability distributions can be based on empirical evidence, expert judgment, or existing practice in the literature.

The corresponding distribution of the endogenous output Y is unknown but can be inferred from solving the model on a large sample of points drawn from the distributions specified for the input parameters—a technique called *uncertainty propagation*. The obtained sample of input–output data can be used to visualize the distribution of the output implied by the specification of uncertainty in $\mathbf{\Theta}$ and to estimate its moments.

3.3.2 Univariate effects functions

One function that explicitly characterizes the relation between a single input parameter Θ_i and the output Y is the so-called *univariate* (or *marginal*) effect function, which represents the conditional expectation of output given the parameter Θ_i,

$$\mathcal{M}^i(\Theta_i) = \mathbb{E}\left[\mathcal{M}(\mathbf{\Theta}|\Theta_i)\right]. \tag{3.14}$$

Focused on a single parameter, its univariate effect function reflects the direction in which Θ_i influences the output, as well as the (non-)linearity and magnitude of the relationship for every possible value of Θ_i. Thanks to these properties, a few recent studies consider univariate effects an informative part of sensitivity analysis, preceding the prioritization of input factors (Younes et al., 2013; Deman et al., 2016; Harenberg et al., 2019).

3.3.3 Variance decomposition framework

Variance-based GSA methods discriminate among all input factors of the model based on their contribution to the variance of its output. The corresponding sensitivity measures derive from Sobol's decomposition framework (Sobol, 1993).

For a square integrable function $g(\mathbf{x})$ over the k-dimensional unit hypercube Ω_k, Sobol's

decomposition reads as

$$g(\mathbf{x}) = g_0 + \sum_{i=1}^{k} g_i(x_i) + \sum_{1\leq i<j\leq k} g_{i,j}(x_i, x_j) + \cdots + g_{1,2,\cdots,k}(x_1, x_2, \cdots, x_k), \tag{3.15}$$

with the summands

$$\begin{cases} g_0 = \mathbb{E}[g(\mathbf{x})] \\ g_i(x_i) = \mathbb{E}[g(\mathbf{x})|x_i] - g_0 \\ g_{i,j}(x_i, x_j) = \mathbb{E}[g(\mathbf{x})|x_i, x_j] - g_0 - g_i(x_i) - g_j(x_j) \\ \cdots \end{cases}. \tag{3.16}$$

Sobol (1993) shows that this expansion is uniquely defined when the summands are *orthogonal* to each other. Applying the representation (3.15) above to our model in (3.13) and given the orthogonality of the summands, the variance of the output, $\text{Var}[Y] = \mathbb{E}\left[(\mathcal{M}(\mathbf{\Theta}) - \mathbb{E}\left[\mathcal{M}(\mathbf{\Theta})\right])^2\right]$, can be decomposed into the partial variances,

$$\text{Var}[Y] = \sum_{i=1}^{M} \text{Var}[\mathcal{M}_i(\Theta_i)] + \sum_{1\leq i<j\leq M} \text{Var}[\mathcal{M}_{i,j}(\Theta_i, \Theta_j)] + \cdots + \text{Var}[\mathcal{M}_{1,2,\cdots,M}(\mathbf{\Theta})]. \tag{3.17}$$

Practically, partial variances measure potential *reductions* in the total variance of the output Y when the values of the corresponding sets of input factors are known. The associated sensitivity measures, *Sobol indices*, are the shares of the partial variances in the total variance of the model. That is, for any subset of parameters' indices $\mathbf{u} \subset \{1, \cdots, M\}$ Sobol indices are defined as

$$S_{\mathbf{u}} = \frac{\text{Var}\left[\mathcal{M}_{\mathbf{u}}(\Theta_{\mathbf{u}})\right]}{\text{Var}[Y]}. \tag{3.18}$$

In particular, the *first order Sobol index* for any parameter Θ_i, defined as

$$S_i = \frac{\text{Var}\left[\mathcal{M}_i(\Theta_i)\right]}{\text{Var}[Y]}, \tag{3.19}$$

reflects the individual impact of this parameter; the second order indices of Θ_i measure the importance of its pairwise interactions; higher order indices the importance of interaction with more parameters in $\mathbf{u}$. The *total Sobol index*,

$$S_i^T = \sum_{i\in\mathbf{u}} S_{\mathbf{u}}, \tag{3.20}$$

quantifies the total contribution of the parameter to the variance in the model.

Sobol indices are appealing measures of sensitivity for a model independently of its nature and complexity. They characterize the effect of its exogenous parameters over the entire

specified uncertainty domain and explicitly quantify individual as well as joint effects. The challenging feature, however, is the associated computational cost—hence the relevance of more efficient methods for computing these indices (Saltelli et al., 2008; Blatman and Sudret, 2010b; Saltelli and D'Hombres, 2010).

3.3.4 Polynomial chaos expansions

The traditional approach to calculating Sobol indices relies on Monte Carlo simulations, known for both their asymptotic optimality and computational demands.[6] The estimation of a single sensitivity index from the decomposition above via Monte Carlo methods typically requires 10^3 to 10^4 model evaluations. In an attempt to reduce the computational burden and make GSA methods applicable to large models, new methods for computing sensitivity indices emerge, with *polynomial chaos expansions* (PCE) claiming special attention.

PCE approximate the outcome of a model with a sum of multivariate orthogonal polynomials of its parameters' values, $\Psi_\alpha(\mathbf{\Theta})$,

$$Y \approx \sum_{\alpha \in \mathcal{A}} y_\alpha \Psi_\alpha(\mathbf{\Theta}). \tag{3.21}$$

The specific form of orthogonal polynomials $\Psi_\alpha(\mathbf{\Theta})$ is implied by the assumed PDFs of parameters and can be derived using the Gram–Schmidt orthogonalization procedure. Among commonly used distributions, for example, the uniform distribution would transfer into the Legendre polynomials; the normal distribution into the Hermite polynomials.[7] For non-standard distributions, random variables can be transformed into ones having a standard distribution (see, for example, Le Gratiet et al., 2017). The approximation (3.21) thus carries the characteristic of the *uncertainty* initially specified for the exogenous inputs of the model.

For practical purposes the set of indices α in (3.21) is truncated (for example, by the maximum degree of the polynomials $\Psi_\alpha(\mathbf{\Theta})$) to obtain a reduced set, $\mathcal{A} \subset \mathbb{N}^M$.

The coefficients of the expansion, $\{y_\alpha\}$, can be estimated via *least-square minimization* over the error ε in

$$Y = \sum_{\alpha \in \mathcal{A}} y_\alpha \Psi_\alpha(\mathbf{\Theta}) + \varepsilon. \tag{3.22}$$

A sample of points for the estimation can be generated by Monte Carlo simulation or other experiment design technique such as Latin hypercube sampling or quasi-random

[6] The reader is referred to Saltelli and D'Hombres (2010) for an overview of Monte Carlo methods for computing sensitivity indices.

[7] See Xiu and Karniadakis (2002) for the correspondence between some families of polynomials and the probability distributions with respect to which they are orthogonal.

sequences (Mckay et al., 1979; Woods and Lewis, 2017).

Applying the standard least-square approach is straightforward yet hardly feasible for large polynomial bases. With a growing number of parameters in $\mathbf{\Theta}$ or elements in $\mathcal{A}$ the curse of dimensionality comes into play. Some recently proposed techniques successfully overcome the issue by benefiting from the fact that typically only a small share of the coefficients $\{y_\alpha\}$ are different from zero. One way to build a parsimonious version of approximation (3.21) is *least-angle regression* (LAR, Efron et al., 2004)—the forward selection algorithm that discriminates the regressors based on their correlation with the output Y. The size of the sparse polynomial basis is in this case limited by a maximum number of elements in (3.21) or by a threshold value of the error of approximation.

The *goodness of least-square minimization* is directly measured by the empirical error

$$\varepsilon_{emp} = \frac{1}{n}\sum_{k=1}^{n}\left(\mathcal{M}(\mathbf{\Theta}_k) - \mathcal{M}^{PC}(\mathbf{\Theta}_k)\right)^2, \tag{3.23}$$

where $\mathbf{\Theta}_k$ is the k-th vector of the values for the parameters, n is the sample size, and $\mathcal{M}^{PC}(\mathbf{\Theta}_k)$ is the PCE of the original model estimated for this vector. This sample-based estimator, however, is subject to over-fitting and cannot be used as a model selection criterion. Cross-validation technique offers an alternative, superior yet cheap estimator (Blatman and Sudret, 2010a). It divides the sample into a training set and a validation set; in the case of *leave-one-out cross-validation* the former contains all but one "point" $\mathbf{\Theta}_k$, and leads to the expansion $\mathcal{M}^{PC\backslash k}(\mathbf{\Theta}_k)$. The corresponding leave-one-out error, ε_{LOO}, is the average squared prediction error across the validation points,

$$\varepsilon_{LOO} = \frac{1}{n}\sum_{k=1}^{n}\left(\mathcal{M}(\mathbf{\Theta}_k) - \mathcal{M}^{PC\backslash k}(\mathbf{\Theta}_k)\right)^2. \tag{3.24}$$

The leave-one-out error is sometimes expressed in relative terms—as its ratio to the total variance of the output. Equation (3.24) can be conveniently reformulated using the full expansion $\mathcal{M}^{PC}(\mathbf{\Theta})$ and the initial experimental design matrix (Blatman, 2009),

$$\varepsilon_{LOO} = \frac{1}{n}\sum_{k=1}^{n}\frac{\left(\mathcal{M}(\mathbf{\Theta}_k) - \mathcal{M}^{PC}(\mathbf{\Theta}_k)\right)^2}{1 - d_k^A}, \tag{3.25}$$

where d_k^A is the k-th diagonal element of the matrix $A(A^TA)^{-1}A^T$, with matrix A containing the values of the basis polynomials at the sample points ($A_{kj} = \Psi_j(\mathbf{\Theta}_k), \quad k = 1, \cdots, n; \quad j = 1, \cdots, card\ \mathcal{A}$). The leave-one-out error can therefore be directly estimated from the PCE performed only once on the original sample.[8]

[8]To account for potential underestimation of the error in small samples without further model evaluations, ε_{LOO} can be multiplied by an additional correction factor; see Chapelle et al. (2002).

An informative statistic based on ε_{LOO} and analogous to the determination coefficient R^2 in the standard regression approach is the so-called *predictivity coefficient*

$$Q^2 = 1 - \frac{\varepsilon_{LOO}}{\frac{1}{n}\sum_{k=1}^{n}\left(\mathcal{M}(\mathbf{\Theta}_k) - \mu_Y\right)^2}, \tag{3.26}$$

where μ_Y is the sample mean of the output Y. The reference values of Q^2 that enable accurate estimation of Sobol indices range from 0.99 to 0.999 (Blatman and Sudret, 2010b; Le Gratiet et al., 2017).

3.3.5 Key advantages offered by PCE

The use of PCE enables several key features. First, the expansion (3.21), denoted by $\mathcal{M}^{PC}(\mathbf{\Theta})$, is itself a ready-to-use *meta-model* (also called a *surrogate model* or *emulator*) of the initial model; the meta-model can substitute the original model for any further simulations, including Monte Carlo simulations for approximating the distribution of the output.

Second, the coefficients of the expansion carry the information about the *first two moments of the output*; namely, its mean, μ_Y^{PC}, and variance, $\mathrm{Var}^{PC}[Y]$, can be estimated from

$$\mu_Y^{PC} = \mathbb{E}\left[\mathcal{M}^{PC}(\mathbf{\Theta})\right] = y_0, \tag{3.27}$$

$$\mathrm{Var}^{PC}[Y] = \mathbb{E}\left[\left(\mathcal{M}^{PC}(\mathbf{\Theta}) - \mu_Y^{PC}\right)^2\right] = \sum_{\alpha\in\mathcal{A},\alpha\neq 0} y_\alpha^2. \tag{3.28}$$

Third, *univariate effects functions* can be expressed in terms of the zero- and single-variable components of the expansion (3.21),

$$\mathbb{E}\left[\mathcal{M}(\mathbf{\Theta}|\Theta_i)\right] = y_0 + \sum_{\alpha\in\mathcal{A}_i} y_\alpha \Psi_\alpha(\Theta_i), \quad \mathcal{A}_i = \{\alpha\in\mathcal{A} : \alpha_i > 0, \alpha_{i\neq j} = 0\}. \tag{3.29}$$

Finally, and most importantly, the *full set of Sobol indices* can be computed analytically, by rearranging the squared coefficients of the expansion (3.21) as follows (Sudret, 2008):

$$S_{\mathbf{u}}^{PC} = \frac{\sum_{\alpha\in\mathcal{A}_\mathbf{u}} y_\alpha^2}{\mathrm{Var}^{PC}[Y]} = \frac{\sum_{\alpha\in\mathcal{A}_\mathbf{u}} y_\alpha^2}{\sum_{\alpha\in\mathcal{A},\alpha\neq 0} y_\alpha^2}. \tag{3.30}$$

The use of PCE therefore enables a broad analysis of uncertainty in the model, including the inference about the distribution of output, the relationship between the output and individual exogenous parameters, and a full decomposition of the variance of the model. An important property of the method is the drastic reduction in the number of

model evaluations required for the analysis, with only the number needed for an accurate approximation being necessary. This property is further explored in Section 3.4.6 below.

To apply the GSA methodology discussed above we use UQLab software framework (Marelli and Sudret, 2014), which is especially designed for uncertainty quantification and sensitivity analysis.

3.4 PCE-based GSA applied to DICE

We first use the PCE-based GSA method to examine the sensitivity of the model with respect to a subset of only few input parameters—in a setting analogous to that in Nordhaus (2008). We likewise choose the *social cost of carbon in the year 2005* as the main output quantity of interest and obtain the results consistent with that earlier study. The inference stays robust to changes in the assumptions of probability distributions specified for the input parameters.

The subsequent analysis of the full set of exogenous parameters, however, records a notable difference in their ranking, featuring ones neglected in the aforementioned earlier study's setting. In such a full setting, in turn, we question the applicability of the assumption of the independence of all the parameters to the DICE model and adjust the GSA setting such that the intrinsic relationships essential for the model's structure are maintained throughout the analysis. As a result, some parameters involved in these relationships appear more significant than the previously top-ranked ones. We conclude the section with a demonstration of computational efficiency of the PCE-based GSA.

3.4.1 Analysis of a subset of parameters

Respecting the choices made by Nordhaus (2008), we preselect the same eight input parameters (also used for the local analysis in Section 3.2.2) that presumably affect the inference from the model the most, and assume that they are normally distributed, as summarized in Table 3.1. We next solve the model on a sample of points generated from these distributions. The low computational cost of DICE enables us to generate a large set of input–output data for the subsequent PCE approximation (3.21). Yet we find that for a PCE with the maximum polynomial degree 3 a sample of size as small as 200 satisfies the predictivity criterion $Q^2 > 0.99$ from (3.26). We therefore use the expansion based on this sample size for further calculations.[9]

Once the PCE approximation is obtained, we can use it as a meta-model and visualize the

[9]Increasing the sample size up to 1000 raises the predictivity coefficient Q^2 to 0.999 but does not noticeably influence the outcome of the sensitivity analysis. See Section 3.4.6 for a demonstration of gains in goodness of approximation for larger samples and higher polynomial degrees.

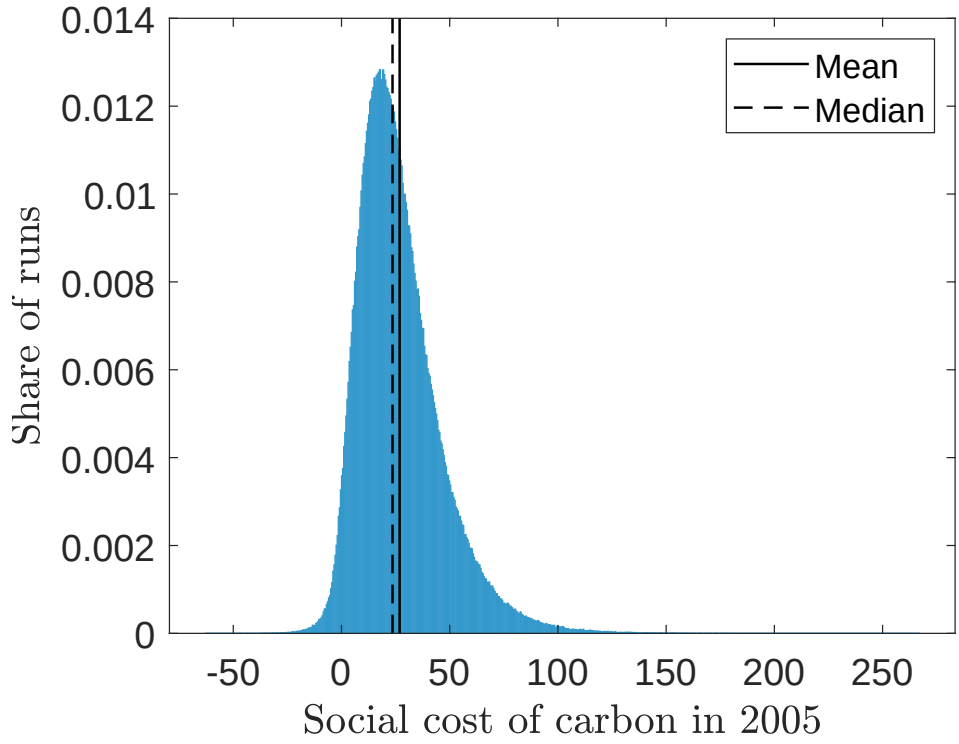


Figure 3.3: Histogram of the output quantity (social cost of carbon in 2005) generated from the PCE meta-model of DICE by Monte Carlo simulations of size 10^6.

distribution of the social cost of carbon via Monte Carlo simulations at almost no cost. Figure 3.3 shows the histogram of the output quantity from 10^6 evaluations of the meta-model: given our uncertainty specification for the input parameters, the distribution of the optimal carbon tax appears right-skewed, with an estimated mean \$26.88 and standard deviation of \$19.15.

The univariate effect functions of the eight parameters computed via equation (3.29) are shown in Figure 3.4 and display a few important features. The slope and nonlinear shape of the function for the initial growth rate of TFP, g_{A0}, suggests that its decline would induce a dramatic increase in the social cost of carbon, while faster technological progress would not imply a proportional decline in the SCC. Both the damage function coefficient, a_2, and the asymptotic population limit, L_{ASYM}, affect the SCC approximately linearly—the former inducing a much stronger marginal increase. The climate sensitivity parameter, λ, has a similar marginal effect at the mean level of the SCC; smaller values of λ are associated with a much lower—even negative for values of λ close to zero—cost of carbon, whereas higher climate sensitivity gradually increases the costs. Changes in the remaining four parameters are not expected to notably affect the optimal carbon tax. From this figure alone the first-order Sobol indices are expected to be high for a_2, λ, and g_{A0}. Total Sobol indices are likely to follow the gradation, unless the interaction effects dominate the individual effects.

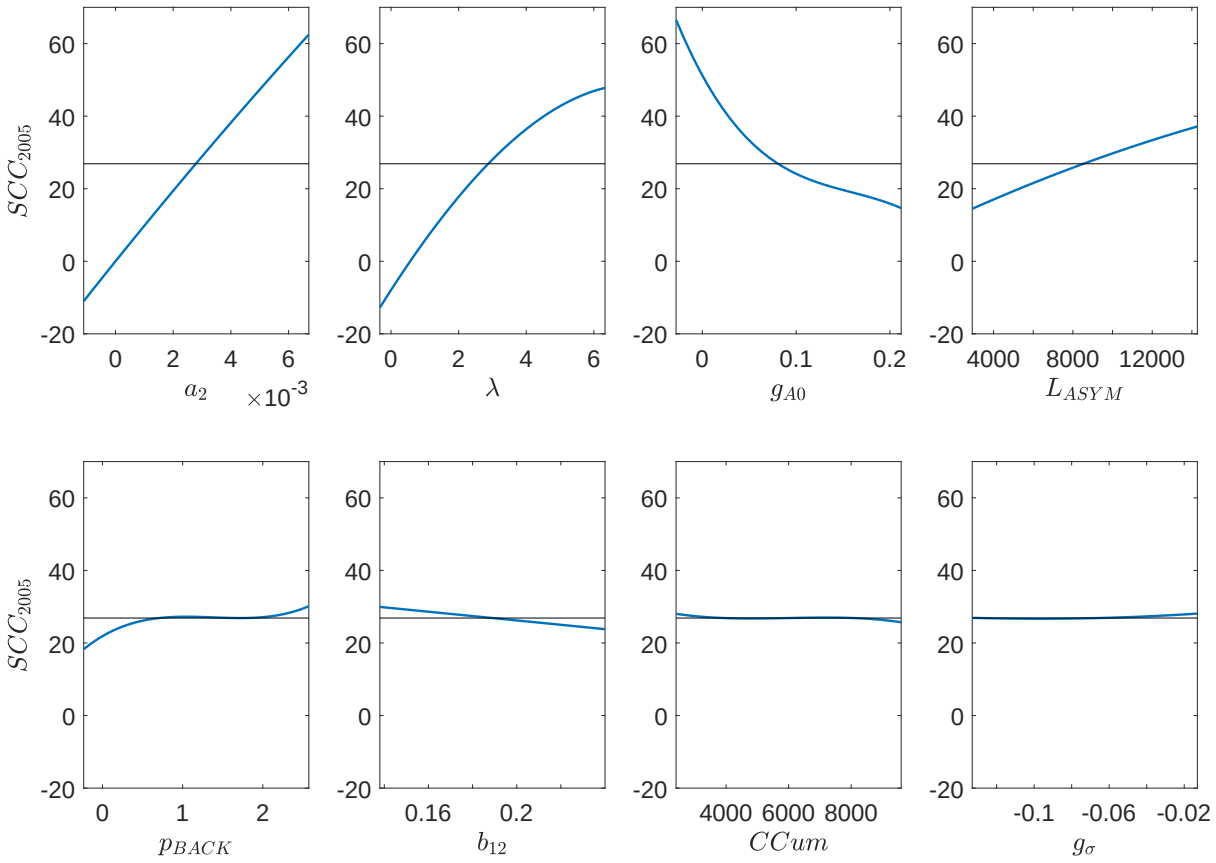


Figure 3.4: Univariate effects of the eight selected parameters sorted by the magnitude of their marginal effects. The horizontal black lines indicate the mean value of the output variable SCC_{2005}.

The first and total Sobol indices computed from (3.30) are shown in the left panel of Figure 3.5. The results are consistent with Nordhaus (2008) and with our expectation: The damage function coefficient, a_2, contributes to the variance of the output the most, followed by climate sensitivity, λ, and the initial growth rate of TFP g_{A0}. Asymptotic population size, L_{ASYM}, plays a far less significant role, while the effect of the uncertainty in the rest of the parameters is negligible.

3.4.2 Effect of a change in parameter distributions

Parameter uncertainty in IAMs is highly debated but only partially explored in the literature: expert opinions on the importance of particular inputs (let alone their probability distributions) often diverge, hampering policy assessment and the decision-making process (Tol, 1995; Heal and Millner, 2014; Athanassoglou, 2015). To address the potential concern about the strong dependence of the results of the analysis on the specification of PDFs, we examine the effect of varying these distributional assumptions on the ranking of the sensitivity indices.

In particular, we compare the benchmark case of normality assumed above to the most

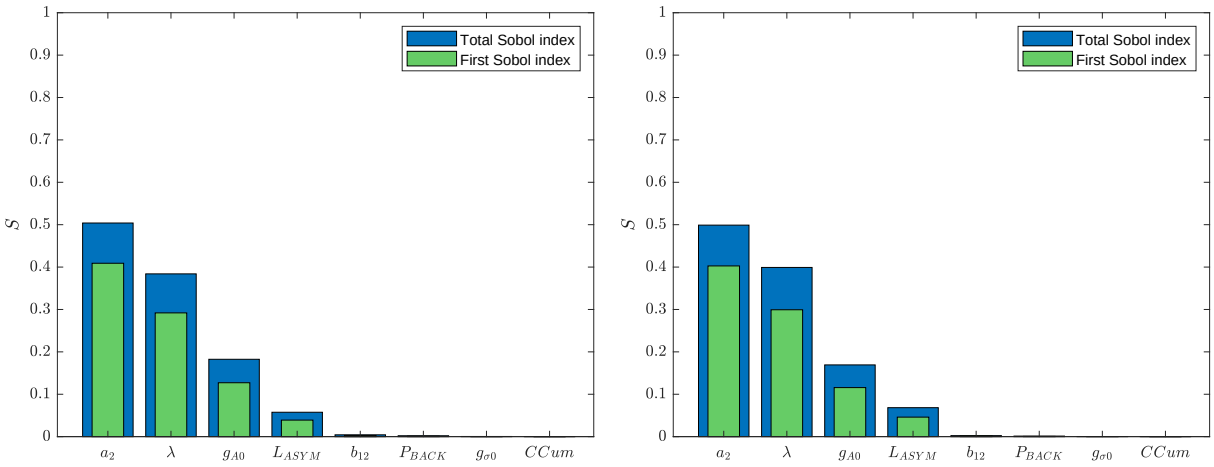


Figure 3.5: First and total Sobol indices for the eight selected parameters given the assumptions of normal (left) vs uniform (right) distributions.

general case—that of uniform distributions. For each of the eight analyzed parameters, we specify a uniform distribution bounded by ±2 standard deviations of their corresponding normal distributions.[10] The right panel of Figure 3.5 demonstrates that the variance decomposition among the parameters is not noticeably affected by switching to the alternative specification.

3.4.3 Analysis of the full set of parameters

The inference above relies on the preliminary subjective selection of a set of parameters that potentially have a strong effect on the SCC. In this section we take full advantage of the global method and expand the analysis to the complete set of *all parameters* of DICE to check if the eight selected parameters are indeed the most influential ones.

Only a very few of the parameters of DICE have their distributions discussed in the literature; hence, it would be highly speculative to impose specific assumptions on the PDFs of all parameters. We therefore follow the practice of Harenberg et al. (2019) and Anderson et al. (2014) and perform a generalized experiment by treating all parameters equally and assuming that we do not have any information on their probability distributions. We assign *uniform* distribution as one maximizing information variability to all parameters, with the support defined by $[0.9 \times \bar{\Theta}_i, 1.1 \times \bar{\Theta}_i]$, where $\bar{\Theta}_i$ is the default value of the parameter Θ_i in DICE.

The approximation (3.21) built on the full set of 36 input parameters naturally requires a notably larger sample size and a reasonable restriction to be imposed on the degree of the multivariate polynomials, so that the estimation of the coefficients, $\{y_\alpha\}$, stays feasible. We choose to restrict the total polynomial degree to 2 and find that a sample of size 500

[10] The results of the sensitivity analysis shown here also hold for a wider support (±3 standard deviations of the corresponding normal distributions) of the uniform distributions.

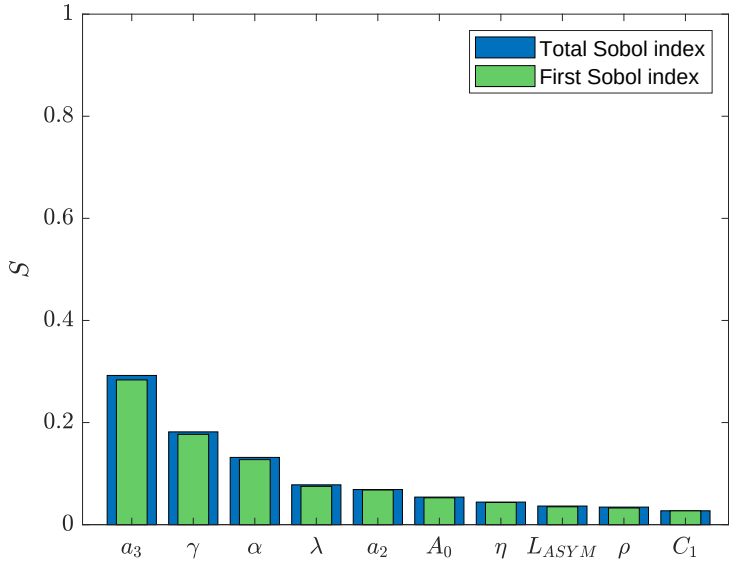


Figure 3.6: First and total Sobol indices for the 10 most important out of all parameters, for the *social cost of carbon in 2005* as the output quantity of interest.

is enough to reach a predictivity level as high as $Q^2 = 0.994$.

The computed Sobol indices reveal a notable difference to the previous, restricted setting. Figure 3.6 shows the parameters with the 10 largest total Sobol indices (the parameters are listed in Table 3.3). The biggest effects are attributed to the damage function exponent, a_3, capital elasticity, γ, and elasticity of marginal utility of consumption, α—all three missing in the subset selected for the previous analysis. The resulting ranking therefore suggests that it can be highly misleading to rely on a *prior* subjective judgment about the importance of input parameters. To infer a model's sensitivity to the uncertainty in its input factors it is necessary to consider their complete set and discriminate among the parameters based on a *full* quantitative decomposition of uncertainty in the model.

The ranking and the conclusion above is in agreement with Anderson et al. (2014), even though we do not find significant interaction effects: all second-order interactions together stand for just over 3% of the total variance in the output.

3.4.4 The issue of dependence

The sensitivity analysis in the previous sections appears comprehensive yet it overlooks one important issue—the implausibility of the assumption of independence among the parameters of an IAM. In fact, the GSA literature warns against the universal treatment of all parameters as independent. Formally, in the presence of correlations or other dependencies the orthogonality of the summands in Sobol decomposition (3.15) can no longer be ensured, and hence the variance decomposition (3.17) is not straightforward

Table 3.3: Parameters with the highest total Sobol indices and their initial values.

Parameter	**Definition**	**Value**
a_3	Damage function exponent	2
γ	Capital elasticity	0.3
α	Elasticity of marginal utility of consumption	2
λ	Climate sensitivity	3
a_2	Damage function coefficient	0.0028
A_0	Initial level of TFP	0.027
η	Radiative forcing of equilibrium CO_2 doubling	3.8
β	Discount factor	0.05
L_{ASYM}	Asymptotic population	8,600
C_1	Temperature transition parameter	0.22

to obtain (Borgonovo and Plischke, 2016). Saltelli and Tarantola (2002) show that the spurious significance of the sensitivity indices is one result of such a misapplication of GSA based on variance decomposition. They propose a framework for the analysis of models with correlated inputs and advocate the use of first-order variances (or, in relative terms, first-order sensitivity indices) as sensitivity measures in the presence of correlations.

In the case of DICE, the issue is even stronger than some intrinsic correlation between its parameters. Not only do some parameters depend on one another, but such dependences are implicitly embedded into the model as its *fundamental relationships*. William Nordhaus emphasizes one such fundamental relationship—that between discount rate and consumption elasticity—in his critical response to the Stern Review (Stern et al., 2006),

> *However, in calibrating a growth model, the time discount rate and the consumption elasticity cannot be chosen independently if the model is designed to match observable real interest rates and savings rates.* Nordhaus (2008, p. 178)

The fallacy of by default imposing the assumption of independence on all parameters also enters GSA analysis: a model is often considered as a "black box" when its structure appears relatively complex. In their global analysis of DICE, both Anderson et al. (2014) and Butler et al. (2014), for example, assume the independence of the parameters and pool them all together. As the rest of this section demonstrates, careful practice leads to a different inference on the factors that determine the value for SCC in the model.

We identify several relationships that are critical to maintain within the structure of DICE and analogous IAMs under GSA scrutiny.

Ramsey rule. In its nature, DICE is a neoclassical optimal growth model that maximizes social welfare in the utilitarian framework. One well-established result is that, in such a framework, when the social optimum is achieved, a basic economic relation called the

Ramsey rule has to hold. The Ramsey rule states that, in an equilibrium, the optimal rate of return on capital r^* is defined by the discount rate ρ, the consumption elasticity α, and the equilibrium growth rate of consumption g^* as

$$r^* = \rho + \alpha g^*. \tag{3.31}$$

In contrast to the original setting of the optimal growth theory, in DICE both the growth rate, g, and the rate of return on capital, r, are endogenous, and for any pair $\{r^*, g^*\}$ there are multiple pairs $\{\rho, \alpha\}$ that would lead to a solution close to these optimal values. The value of consumption elasticity, α, is therefore calibrated *in conjunction with* the pure rate of time preference, ρ. Consequently, every time the assumption on α changes the model has to be recalibrated such that the new value of ρ is the one that brings the optimal rate of return on capital the closest to the benchmark value of 0.055 (Nordhaus, 2008, p. 61).

For our analysis we therefore consider the pair $\{r^*, \alpha\}$—not $\{\rho, \alpha\}$—as independent. Our GSA algorithm randomly generates the values for α and target values for r around their central values. The rate of time preference, ρ, is consequently calibrated such that it brings the return on capital, r, the closest to the target level r^*.

Climate damages. The damage function in an IAM quantifies the impact of the increase in atmospheric temperature on the economic output. In DICE this function takes the reduced quadratic form

$$\Omega_t = \frac{1}{1 + a_2 T_{A_t}^{a_3}} \tag{3.32}$$

calibrated to match the estimated damage to the world GDP, aggregated from 12 regions, at the levels of warming 0℃, 2.5℃, and 6℃ (Nordhaus and Boyer, 2000). Because the values for the parameters a_2 and a_3 are set simultaneously to match the target levels of damage, we need to adjust their sampling. Following van den Bijgaart et al. (2016), we define ω as the level of damage associated with the benchmark level of 3℃ and then for every randomly chosen value of a_3 set $a_2 = 3^{-a_3}\omega$. Such reformulation allows us to pick the values for both ω and a_3 independently.

Production function. The production function in DICE is of a standard Cobb–Douglas form with capital K, labor L, and technology A as its inputs and the elasticity of capital γ. As defined in the model, this relationship should hold for the initial stage of the economy as well. This implies that the calibrated parameters A_0 and γ cannot be simultaneously set to arbitrary values. The initial level of technology, A_0, instead has to be set to satisfy

$$Y_0^{GROSS} = A_0 K_0^{\gamma} L_0^{1-\gamma}. \tag{3.33}$$

For this reason, every time an algorithm randomly picks a value for γ and initial values

for Y_0, K_0, and L_0 from the corresponding distributions, A_0 is determined by these values.

Initial state of the climate system. One physical property of the climate system is that the warming induced by a doubling in the concentrations of CO_2 stays unchanged (and this law holds for most of greenhouse gases); that is, the average global surface temperature rises by the same amount each time the concentration increases twofold. A key notion that characterizes the climate module of an IAM is thus its *equilibrium climate sensitivity*, $\Delta T_{2\times CO_2}$. This parameter defines the expected rise in the average global surface temperature, T_{A_t}, in response to a doubling in the concentrations of CO_2, M_{AT_t}, with respect to some reference level at time $t = 0$, typically taken as the preindustrial level (Flato et al., 2013). The change in atmospheric temperature induced by any change in CO_2 concentrations can therefore be inferred from the number of such "doublings". Formally, the long-term equilibrated temperature change can thus be approximated by the following relationship:

$$T_{A_t} = \Delta T_{2\times CO_2} \frac{\log\left(\frac{M_{AT_t}}{M_{AT_0}}\right)}{\log(2)}, \tag{3.34}$$

where the reference value of concentrations is the preindustrial level, $M_{AT_0} = M_{AT_{1750}}$.

In DICE, the initial level of CO_2 concentrations, $M_{AT_{2005}}$, the initial temperature anomaly, $T_{A_{2005}}$, and climate sensitivity, $\Delta T_{2\times CO_2}$, are the three parameters set to respect this fundamental relationship and therefore should not be disentangled. A logical way to proceed is to pick random values for climate sensitivity and the initial level of atmospheric CO_2 concentrations and infer the temperature anomaly from these two values using (3.34). In this case, however, the two parameters $\Delta T_{2\times CO_2}$ and $M_{AT_{2005}}$ cannot be drawn independently: the lower bound of the distribution of $M_{AT_{2005}}$ depends on the value of $\Delta T_{2\times CO_2}$ and should be set such that, in accordance with the model's assumptions, the initial temperature anomaly does not drop below 0. Instead, we set an opposite relationship—the algorithm randomly picks $T_{A_{2005}}$ and $\Delta T_{2\times CO_2}$ and calculates $M_{AT_{2005}}$ from (3.34) as[11]

$$\begin{aligned} M_{AT_{2005}} &= \exp\left(\log(M_{AT_{1750}}) + \frac{\log(2)}{\Delta T_{2\times CO_2}} T^{1750}_{A_{2005}}\right) \\ &= \exp\left(\log(M_{AT_{1750}}) + \frac{\log(2)}{\Delta T_{2\times CO_2}} \left(T^{1900}_{A_{2005}} + 0.5883\right)\right). \end{aligned} \tag{3.35}$$

The reformulation thereby leaves us with two independently drawn parameters and the third one, $M_{AT_{2005}}$, determined by their values.

[11] DICE specifies the temperature anomaly with respect to the year 1900, denoted here as $T^{1900}_{A_t}$, whereas the ratio of the concentrations refers to the preindustrial level (the year 1750). Therefore, for consistency of (3.34), we calculate $T^{1750}_{A_t}$ from the initial setting of the model and find the difference $T^{1750}_{A_t} - T^{1900}_{A_t} = 0.5883$.

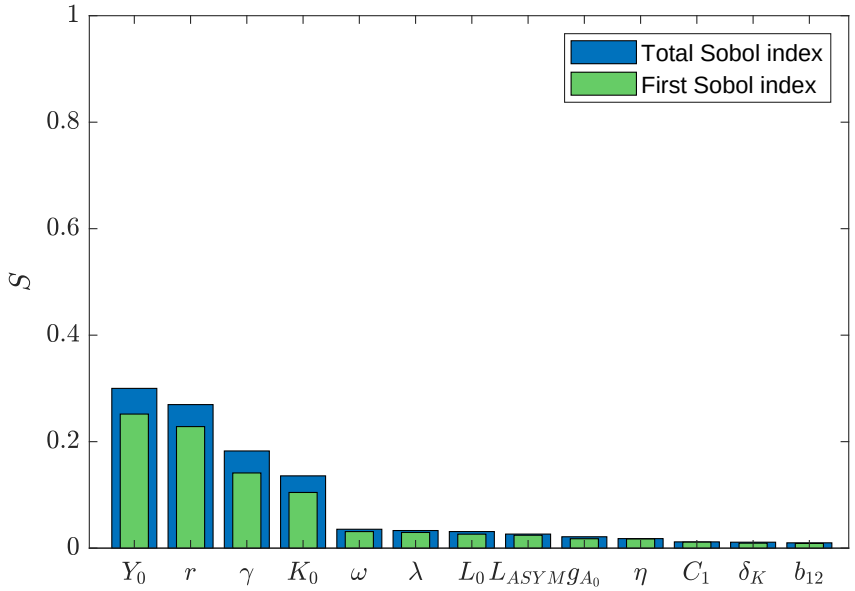


Figure 3.7: First and total Sobol indices for the parameters that explain at least 1 percent of the variance in the SCC, in the setting adjusted for the fundamental relationships.

Short-term adjustment of the climate system. The carbon cycle in DICE can be written in terms of the vector of the atmospheric, upper layer, and lower ocean CO_2 concentrations, $\mathbf{M_t} = [M_{AT_t}\ M_{U_t}\ M_{L_t}]'$, and a transition matrix $\mathbf{B}$ among these layers,

$$\mathbf{M_{t+1}} = \mathbf{B}\mathbf{M_t} + [E_t\ 0\ 0]', \tag{3.36}$$

where

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{21} & 0 \\ b_{12} & b_{22} & b_{32} \\ 0 & b_{23} & b_{33} \end{bmatrix} = \begin{bmatrix} 1-b_{12} & \frac{b_{12}\overline{M_{AT}}}{\overline{M_U}} & 0 \\ b_{12} & 1-b_{21}-b_{23} & \frac{b_{23}\overline{M_U}}{\overline{M_L}} \\ 0 & b_{23} & b_{33} \end{bmatrix}. \tag{3.37}$$

The transition matrix $\mathbf{B}$ is fully described by its two parameters b_{12} and b_{23}, which are calibrated simultaneously to fit a benchmark *impulse response function*. The impulse response function defines the fraction of CO_2 remaining in the atmosphere several decades after an impulse of CO_2 emissions. This target fraction itself can be varied independently, while the pair $\{b_{12}, b_{23}\}$ has to be tuned to bring the system to this new target impulse response.

3.4.5 Final results of the analysis

The largest sensitivity indices from the analysis with all the adjustments are shown in Figure 3.7. Additionally, Appendix 3.B includes the histogram of the simulated value for SCC and reports the strongest univariate effects; the full set of Sobol indices for this

setting is reported in Table 3.B.[12] The most prominent result of the refined setting is that the value for SCC is indeed strongly influenced by the output—in agreement with the *proportionality* claimed in the analytical literature. This and other components of the production function of the same period together explain more than a half of the variance in the value for the optimal carbon tax.

Among all *economic* parameters the most important ones are the same-period (in this case, initial) output and capital levels, Y_0 and K_0, the rates of return on capital, r, and of capital depreciation, δ_K, the capital elasticity, γ, the initial and asymptotic sizes of population, L_0 and L_{ASYM}, and the rate of technological progress, g_{A_0}. None of these parameters directly participates in the analytical formulas 3.12. Two of them—L_0 and L_{ASYM}—enter the formula through the growth rate of population, which in DICE is largely determined by the initial and the asymptotic values. The rate of return on capital, r, can be considered as the parameter that defines the values for two other important parameters—the time discount rate and the elasticity of marginal utility of consumption—and thereby reflects their importance. Another economic factor in the formula—the rate of growth for consumption—is *endogenous* in DICE and stems from the decision variable C_t. A more detailed analysis reveals that the parameters mentioned above, in fact, explain this rate of growth to a large extent. In support of the claim, Table 3.C summarizes the results of a linear regression of the growth rate of consumption on the involved parameters—with Y_0, K_0, r, g_{A_0}, and δ_K placed among the most significant ones. Consistent with this hypothesis, the first three of these parameters exhibit the highest degrees of interaction (see the largest interaction indices reported in Figure 3.C).

The rest of the results are also in line with the analytical inference: Among the parameters of the *climate* module the analysis features those that govern the persistence of atmospheric carbon dioxide, b_{12}; the long-term equilibration of the atmospheric temperature, η and C_1; and the connection of the two—climate sensitivity parameter, λ. Finally, the *interaction* of the two systems is described by the shape of the damage function, and quantified by the severity of the damage to the economy with every degree of temperature rise, ω, which thereby contributes to the marginal costs of CO_2 emissions.

In general, the results of our analysis are consistent with the analytical evidence in the literature and carry the careful dependence-adjusted reformulations of the relationships embedded in DICE.

[12] In this setting adjusted for the fundamental relationships the same sample size—500 model funs—reaches a predictivity level $Q^2 = 0.9846$. Larger samples ($\geq$ 1500 runs) raise the criterion to the recommended benchmark value of 0.99, but they support the same inference—with only minor changes (of order 10^{-3}) in the values for the sensitivity indices and no changes in their ranking.

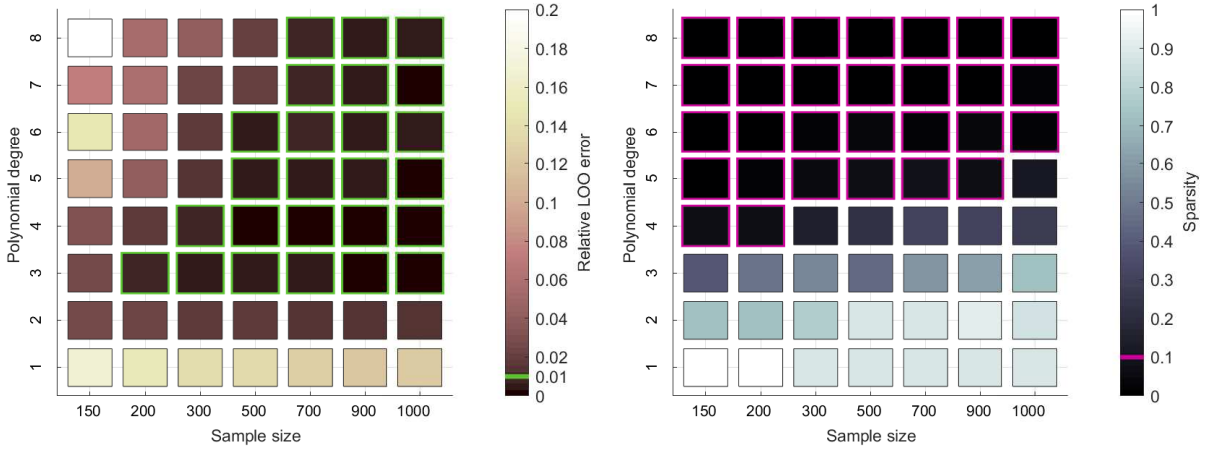


Figure 3.8: Computational efficiency of sparse PCE for the analysis of a subset parameters in Section 3.4.1. The left panel shows relative approximation error for increasing sample sizes and polynomial degrees; the right panel, the sparsity of the corresponding polynomial bases (i.e. the ratio of sparse bases to the full bases).

3.4.6 Computational efficiency

The IAM literature is indeed departing from conventional Monte Carlo methods in search of more efficient ones. In one of the first sensitivity analyses of DICE, Nordhaus (2008) uses 100 model runs to infer the impact of the uncertainty in selected parameters—a relatively expensive assessment of the effect of changes on a restricted parameter space. Turning to global methods and the large samples required, Butler et al. (2014) examine the convergence of individual Sobol indices for the parameters in DICE for the sample sizes of the order 10^6. Anderson et al. (2014) apply a much more efficient method to calculate the sensitivity measures for all parameters; the method still requires a Monte Carlo sample of 10^5 model runs.

The prominent feature of our analysis is the small number of model runs required by the undertaken GSA approach. The efficiency of PCE for sensitivity analysis is demonstrated in Sudret (2008) and Blatman and Sudret (2010b), its fast convergence supported by later practical implementations (e.g., Arwade et al., 2010; Younes et al., 2013; Harenberg et al., 2019).

The computational costs of our *PCE-based* global analysis of DICE are notably low. To demonstrate this argument, the left panel of Figure 3.8 depicts the relative approximation error in the case of a subset of eight parameters in Section 3.4.1, for increasing sample sizes and polynomial degrees. The green squares in this figure mark all settings that meet the goodness of approximation criterion from Section 3.3.4. Additionally, the right panel of this figure shows the corresponding sparsity index of the sparse polynomial basis defined as the ratio of non-zero polynomial coefficients in the approximation (3.21) to the size of the full basis $\mathcal{A}$. For illustration purposes, the pink squares mark the combinations

where less than 10% of the elements are kept in the expansion.

In the analysis in Section 3.4.1 a sample of a size as small as 200 and the maximum polynomial degree of 3 provided an approximation with an error of less than 1% of the total variance. The sparse polynomial basis in this case is less than a half of the full basis. The analysis in Section 3.4.5 is based on the full parameter space and naturally requires a larger sample for constructing an accurate PCE approximation; yet, a sample as small as 500 runs proves sufficient in this case. The PCE-based approach can thereby make global sensitivity analysis an affordable routine for IAM models of higher complexity.

3.5 Conclusion

This paper addresses the call in the literature for more transparency with regard to and a greater scrutiny of IAMs. We argue for incorporating global sensitivity analysis as a common practice in integrated assessment modeling. GSA offers a detailed representation of uncertainty in a model given the uncertainty assumed for its parameters. In particular, variance-based GSA differentiates all exogenous parameters according to their contribution to the total variance of output, made directly or through interrelation with other parameters. Variance-based GSA can be efficiently carried out through polynomial chaos expansions, which substitute the original model with its polynomial approximation.

PCE method provides a "shortcut" to decomposing the variance of the model and informs the modeler about the distribution of the output quantity implied by the initial distributional assumptions on the inputs. The efficient calculation of sensitivity indices via PCE significantly reduces the computational costs of the analysis and thereby enables its application to much more sophisticated IAMs. The sample size in this case needs only to be large enough to enable accurate approximation.

To demonstrate the benefits of the suggested approach, we apply the PCE-based GSA method to a benchmark integrated assessment model, DICE. Our analysis, in agreement with existing studies, suggests that incorporating the full set of its parameters is essential for a viable analysis. An equally important requirement is careful consideration of the relationships among the parameters that are fundamental to the model's structure. Reformulations or recalibrations might be needed for the GSA methods to be applicable.

Given the limited number of model runs required by the implemented efficient GSA method, it renders affordable comprehensive sensitivity analysis for much more complex IAMs and allows them to be included in GSA intercomparison projects.

3.A Parameters in DICE

Overall DICE has 51 exogenous parameters that fully specify its economic module, climate module, the interaction of the two, and the general model settings. The parameters, their full names and default values are listed in Table 3.A below.

Some of the values specified as parameters, however, appear redundant. Namely, the value of the initial level of output, *Q0*, is never used in the code, as the actual initial level is defined by the general specification of the production function for all periods.

Next, the transition matrix for the CO_2 concentrations is in fact determined by two out of its seven coefficients, which makes specifying the values for *b11*, *b21*, *b22*, *b32*, and *b33* unnecessary.

We also chose to omit the parameters that have the default value 0 from the analysis in order to adhere to the initial specification of the model; for example, to the reduced quadratic form of the damage function. These omitted parameters are *DSIG2*, *A1*, and *DPARTFRACT*.

We do not vary the parameters that tune DICE to particular participation or mitigation scenarios. Thus, the maximum rate of mitigation, *LIMMIU*, and the participation parameters *PARTFRACT1*, *PARTFRACT2*, and *PARTFRACT21* are always kept at 1, which enables the unrestricted optimization scenario.

Finally, the two scaling parameters, *scale1* and *scale2*, magnify the utility function but do not influence the calculations and therefore are omitted from the analysis.

All in all, we are left with the set of 36 parameters, and in our analysis refer to it as the full set of exogenous parameters in DICE.

Table 3.A: Parameters in DICE. The parameters that do not participate in this study for the reasons described in Section 3.A are marked gray.

Parameter	Notation	Definition	Value
Preferences			
B_ELASMU	α	Elasticity of marginal utility of consumption	2.0
B_PRSTP	ρ	Rate of social time preference, per year	0.015
Population and technology			
POP0	L_0	Initial population (2005, millions)	6514
GPOP0	g_{L_0}	Growth rate of population, per decade	0.35
POPASYM	L_{ASYM}	Asymptotic population (millions)	8600
A0	A_0	Initial level of total factor productivity	0.02722
GA0	g_{A_0}	Initial growth rate for technology, per decade	0.092
DELA	δ_A	Decline rate of technology change, per decade	0.001
DK	δ_K	Depreciation rate on capital, per year	0.100
GAMA	γ	Capital elasticity in production function	0.300
Q0	Y_0	Initial world gross output (2005, trill US dollars)	61.1
K0	K_0	Initial value of capital (2005, trill US dollars)	137
Emissions and carbon cycle			
SIG0	σ_0	CO_2-equivalent emissions-to-GNP ratio, 2005	0.13418
GSIGMA	g_σ	Initial growth of decarbonization, per decade	-0.0730
DSIG	σ	Decline rate of decarbonization, per decade	0.003
DSIG2	σ_2	Quadratic term in decarbonization	0
ELAND0	E_{TREE_0}	Initial carbon emissions from land (2005, GtC)	11.000
FOSSLIM	$CCum$	Maximum cumulative extraction fossil fuels	6000
MAT2000	$M_{AT_{2005}}$	Initial concentration in atmosphere (2005, GtC)	808.9
MU2000	$M_{U_{2005}}$	Initial concentration in upper strata (2005, GtC)	1255
ML2000	$M_{L_{2005}}$	Initial concentration in lower strata (2005, GtC)	18365
b11	b_{11}	Carbon cycle transition, persistence in atmosphere	0.810712
b12	b_{12}	Carbon cycle transition, atmosphere to upper layer	0.189288
b21	b_{21}	Carbon cycle transition, upper layer to atmosphere	0.097213
b22	b_{22}	Carbon cycle transition, persistence in upper layer	0.852787
b23	b_{23}	Carbon cycle transition, upper to lower layer	0.05
b32	b_{32}	Carbon cycle transition, lower to upper layer	0.003119
b33	b_{33}	Carbon cycle transition, persistence in lower layer	0.996881
Climate model			
T2XCO2	λ	Equilibrium temperature impact of CO_2 doubling (℃)	3
FEX0	F_{EX_0}	Initial forcings of non-CO_2 GHG, 2005	-0.06
FEX1	F_{EX_1}	Estimate of 2100 forcings of non-CO_2 GHG	0.30
TOCEAN0	T_{O_0}	Initial lower ocean temperature change (1900 to 2005, ℃)	0.0068
TATM0	T_{A_0}	Initial atmospheric temperature change (1900 to 2005, ℃)	0.7307
C1	C_1	Climate-equation coefficient for upper stratum	0.220
C3	C_3	Transfer coefficient upper to lower stratum	0.300
C4	C_4	Transfer coefficient for lower stratum	0.050
FCO22X	η	Radiative forcings from equilibrium CO_2 doubling	3.8
Climate damage and abatement			
A1	a_1	Damage function, intercept	0
A2	a_2	Damage function, quadratic term	0.0028388
A3	a_3	Damage function, exponent	2.00
EXPCOST2	θ_2	Exponent of control cost function	2.8
PBACK	p_{BACK}	Cost of backstop, (thousand US dollars per tC)	1.17
BACKRAT	r_{BACK}	Ratio initial to final backstop cost	2
GBACK	$g_{p_{BACK0}}$	Initial cost decline backstop (percent per decade)	0.05
LIMMIU	μ_{MAX}	Upper limit on emissions control rate	1
Participation and other parameters			
PARTFRACT1	f_{PART_1}	Fraction of emissions under control regime in 2005	1
PARTFRACT2	f_{PART_2}	Fraction of emissions under control regime in 2015	1
PARTFRACT21	$f_{PART_{21}}$	Fraction of emissions under control regime in 2205	1
DPARTFRACT	$\delta_{f_{PART}}$	Decline rate of participation	0
scale1	s_1	Scaling coefficient in the objective function	194
scale2	s_2	Scaling coefficient in the objective function	381800

3.B Detailed results of the analysis

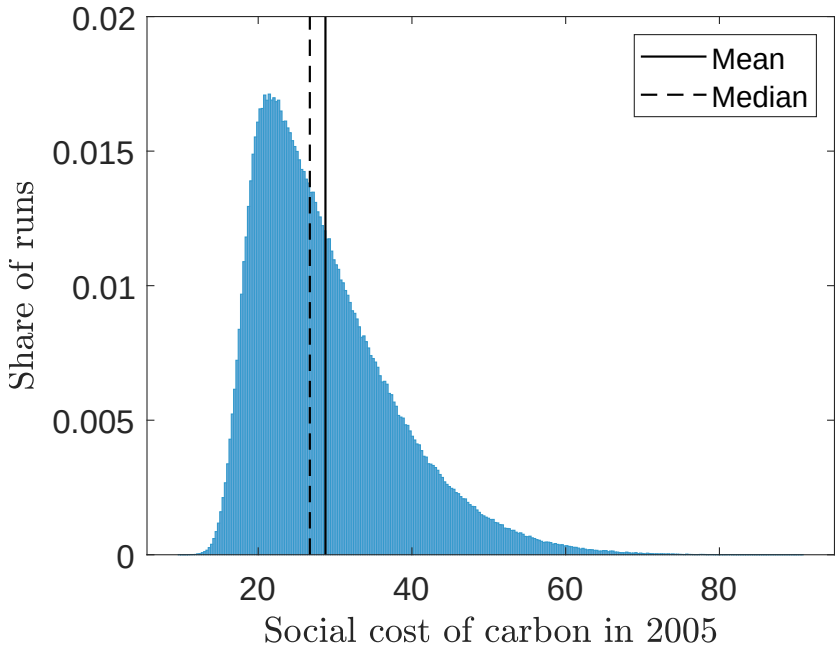


Figure 3.A: Histogram of the output quantity (social cost of carbon in 2005) generated from the PCE meta-model of DICE by Monte Carlo simulations of size 10^6, in the full setting adjusted for the fundamental relationships.

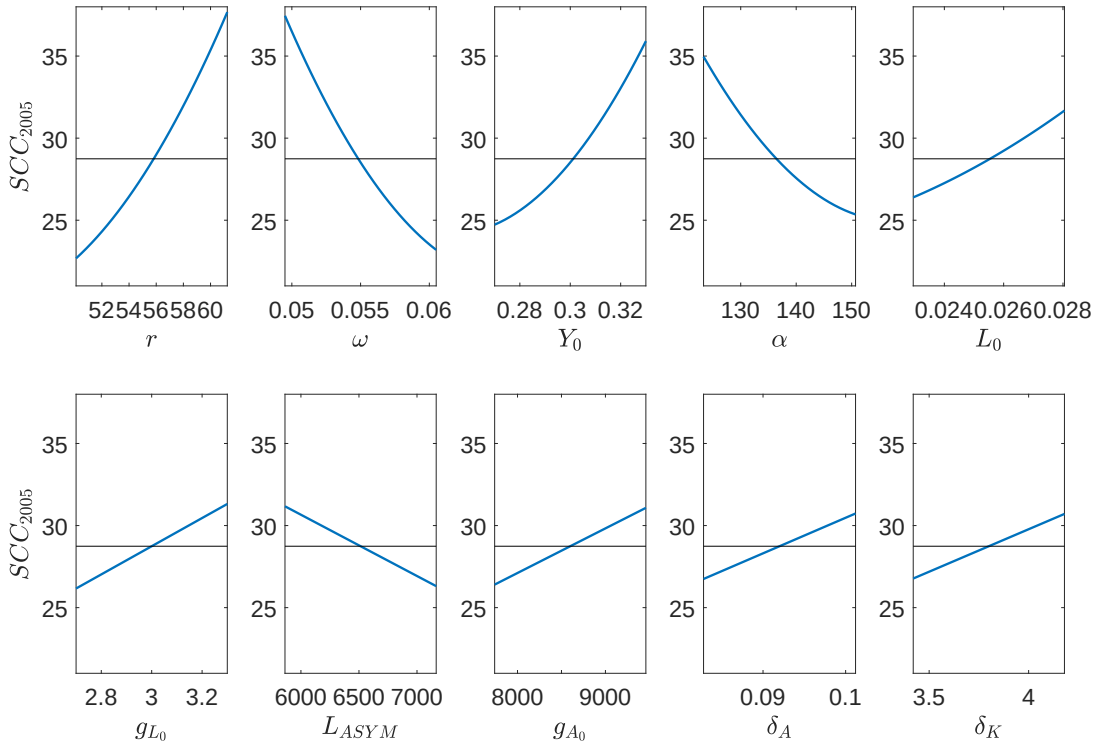


Figure 3.B: The first 10 strongest univariate effects, in the full setting adjusted for the fundamental relationships. The horizontal black lines indicate the mean value of the output variable SCC_{2005}.

Table 3.B: Sobol indices for the full set of parameters in the final, dependence-adjusted setting in Section 3.4.4. For every parameter, the columns S^T and S^1 report its total and first Sobol indices; S^I reports the sum of all interactions of this parameter.

Parameter	**Short definition**	S^T	S^1	S^I
Y_0	Initial output	0.3001	0.2518	0.0482
r	Return on capital	0.2696	0.2283	0.0413
γ	Capital elasticity	0.1825	0.1412	0.0413
K_0	Initial capital	0.1357	0.1046	0.0312
ω	Damage at 1℃	0.0355	0.0311	0.0044
λ	Climate sensitivity	0.0331	0.0295	0.0036
L_0	Initial population	0.0312	0.0263	0.0049
L_{ASYM}	Asymptotic population	0.0264	0.0243	0.0021
g_{A_0}	Initial growth rate for TFP	0.0214	0.0177	0.0037
η	Radiative forcings from CO_2 doubling	0.0179	0.0172	0.0007
C_1	Climate coefficient, upper stratum	0.0117	0.0113	0.0005
δ_K	Depreciation rate on capital	0.0112	0.0094	0.0018
b_{12}	Carbon transition, atm. to upper layer	0.0100	0.0088	0.0012
$M_{U_{2005}}$	Initial concentration in upper strata	0.0011	0.0010	0.0001
C_3	Heat transfer upper to lower stratum	0.0011	0.0009	0.0002
a_3	Damage function exponent	0.0009	0.0009	0.0000
T_{A_0}	Initial atmospheric temperature	0.0008	0.0007	0.0001
σ_0	Emissions-to-GNP ratio	0.0007	0.0007	0.0000
σ	Decline rate of decarbonization	0.0005	0.0000	0.0005
θ_2	Exponent of control cost function	0.0003	0.0002	0.0001
p_{BACK}	Cost of backstop	0.0003	0.0001	0.0002
g_{L_0}	Growth rate of population	0.0003	0.0002	0.0001
g_σ	Initial growth of decarbonization	0.0003	0.0000	0.0003
$M_{L_{2005}}$	Initial concentration in lower strata	0.0003	0.0001	0.0002
F_{EX_0}	Initial forcings of non-CO_2 GHG	0.0002	0.0000	0.0002
T_{O_0}	Initial lower ocean temperature	0.0002	0.0000	0.0002
F_{EX_1}	Forcings of non-CO_2 GHG in 2100	0.0002	0.0000	0.0002
$CCum$	Maximum cumulative fossil fuels	0.0001	0.0000	0.0001
E_{TREE_0}	Initial emissions from land	0.0001	0.0000	0.0001
C_4	Transfer coefficient for lower stratum	0.0001	0.0000	0.0001
α	Elasticity of marginal utility	0.0000	0.0000	0.0000
$g_{p_{BACK0}}$	Initial cost decline backstop	0.0000	0.0000	0.0000
δ_A	Decline rate of TFP change	0.0000	0.0000	0.0000
r_{BACK}	Ratio initial to final backstop cost	0.0000	0.0000	0.0000

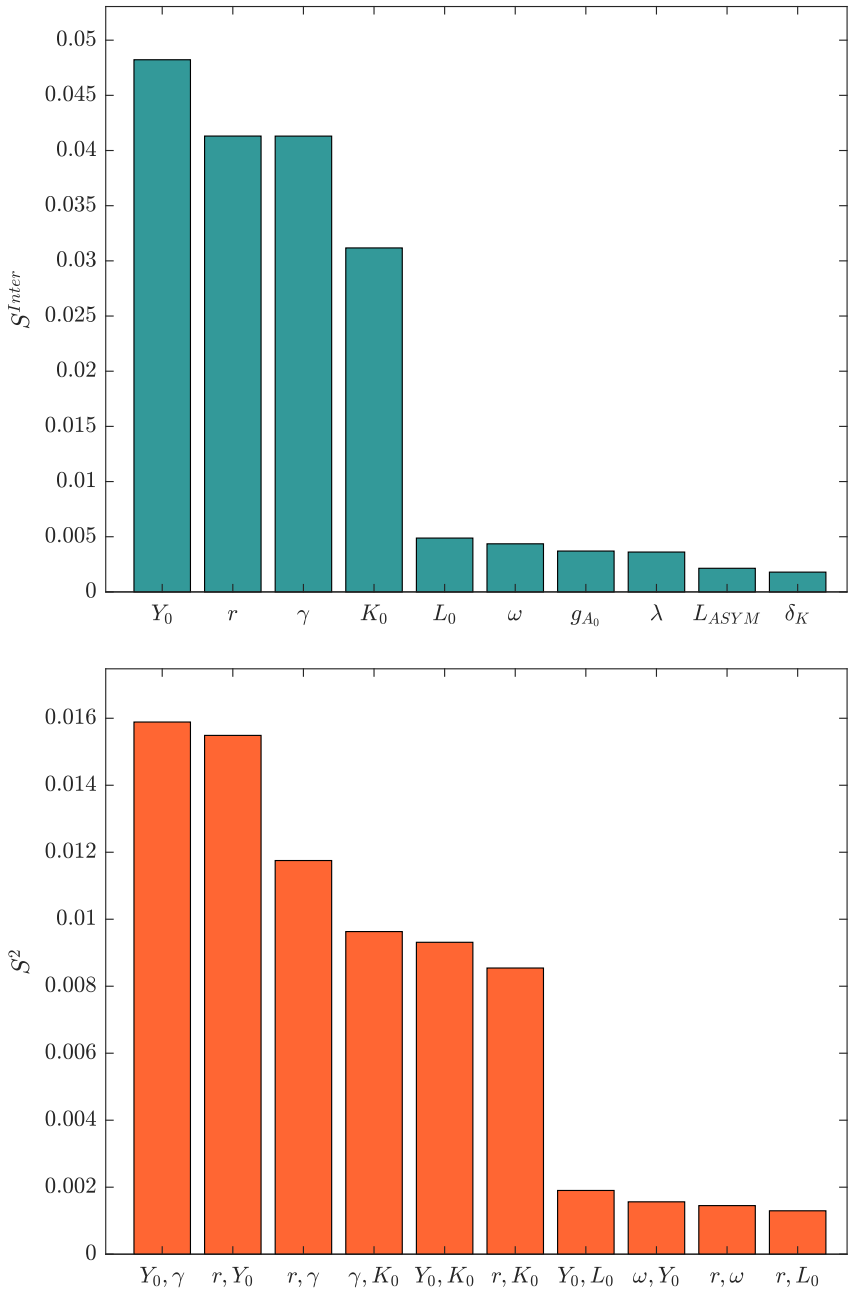


Figure 3.C: The first 10 parameters with the largest second-order Sobol indices (top) and the first 10 largest interactions for pairs of parameters (bottom), in the setting adjusted for the fundamental relationships.

Table 3.C: Results for the regression of the growth rate of consumption in DICE on the input parameters, in the full setting adjusted for the fundamental relationships.

	Dependent variable:			
	Growth rate of consumption			
	Estimate	St. error	t-statistic	p-value
r	-3.0230	0.0211	-143.1783	0
Y_0	0.0064	2.0893e-05	305.1161	0
K_0	-0.0026	8.5296e-06	-305.4387	0
α	-0.0050	0.0006	-8.5828	0
L_0	-6.1192e-06	1.7906e-07	-34.1733	0
L_{ASYM}	4.3936e-06	1.3512e-07	32.5161	0
g_{L_0}	0.0308	0.0033	9.2605	0
g_{A_0}	1.8051	0.0126	143.1294	0
δ_K	-1.1364	0.0116	-97.6792	0
γ	1.4464	0.0039	372.9036	0
η	-0.0009	0.0003	-2.9175	0.0036
ω	-0.1103	0.0455	-2.4221	0.0155
$M_{U_{2005}}$	-2.2244e-06	9.2874e-07	-2.3951	0.0167
b_{12}	0.0073	0.0034	2.1745	0.0298
$CCum$	3.6164e-07	1.9419e-07	1.8623	0.0627
δ_A	-2.0640	1.1642	-1.7729	0.0764
$g_{p_{BACK0}}$	0.0385	0.0232	1.6571	0.0977
T_{A_0}	-0.0023	0.0016	-1.4412	0.1497
C_1	-0.0073	0.0053	-1.3679	0.1715
a_3	0.0007	0.0006	1.2305	0.2187
σ_0	-0.0104	0.0087	-1.1957	0.2320
θ_2	0.0004	0.0004	0.9879	0.3233
C_3	0.0038	0.0039	0.9803	0.3271
T_{O_0}	0.1593	0.1714	0.9290	0.3530
C_4	-0.0207	0.0234	-0.8867	0.3754
F_{EX_0}	0.0164	0.0194	0.8472	0.3970
p_{BACK}	0.0006	0.0010	0.6203	0.5351
λ	-0.0002	0.0004	-0.6072	0.5438
$M_{L_{2005}}$	3.315e-08	6.3264e-08	0.5240	0.6003
g_σ	-0.0079	0.0159	-0.4988	0.6180
E_{TREE_0}	-3.4161e-05	0.0001	-0.3227	0.7470
F_{EX_1}	-0.0012	0.0039	-0.3216	0.7478
σ	-0.0299	0.3870	-0.0773	0.9384
r_{BACK}	-5.3099e-06	0.0005	-0.0091	0.9927
(Intercept)	-0.1508	0.0067	-22.5366	0
Observations	2000			
R^2	0.9948			
Adjusted R^2	0.9947			
Residual Standard Error	0.0030			
F Statistic	1.11e+04			

Note: *p<0.1; **p<0.05; ***p<0.01

Part III

Bibliography and Curriculum Vitae

Bibliography

Alexander, L., Allen, S., Bindoff, N., Breon, F.M., Church, J., Cubasch, U., Emori, S., Forster, P., Friedlingstein, P., Gillett, N., Gregory, J., Hartmann, D., Jansen, E., Kirtman, B., Knutti, R., Kanikicharla, K., Lemke, P., Marotzke, J., Masson-Delmotte, V., Meehl, G., Mokhov, I., Piao, S., Plattner, G.K., Dahe, Q., Ramaswamy, V., Randall, R., Rhein, M., Rojas, M., Sabine, C., Shindell, D., Stocker, T., Talley, L., Vaughan, D., Xie, S.P., 2013. Summary for Policymakers, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press. TF Stocker, D Qin, G-K Plattner, MMB Tignor, SK Allen, J Boschung, A Nauels, Y Xia, V Bex, PM Midgle (ed). pp. 3–30.

Allen, M.R., Frame, D.J., Huntingford, C., Jones, C.D., Lowe, J.A., Meinshausen, M., Meinshausen, N., 2009. Warming caused by cumulative carbon emissions towards the trillionth tonne. Nature 458, 1163–1166.

Anderson, B., Borgonovo, E., Galeotti, M., Roson, R., 2014. Uncertainty in climate change modeling: Can global sensitivity analysis be of help? Risk analysis: an official publication of the Society for Risk Analysis 34, 271–293.

Angrist, J.D., Pischke, J.S., Pischke, J.S., 2013. Mostly harmless econometrics: an empiricists companion. Cram101 Publishing.

Arent, D., Tol, R., Faust, E., Hella, J., Kumar, S., Strzepek, K., Tóth, F., Yan, D., 2014. Key economic sectors and services. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 659-708.

Arwade, S.R., Moradi, M., Louhghalam, A., 2010. Variance decomposition and global sensitivity for structural systems. Engineering Structures 32, 1–10.

Athanassoglou, S., 2015. Multidimensional welfare rankings under weight imprecision: a social choice perspective. Social Choice and Welfare 44, 719–744.

Atkinson, G., Dietz, S., Helgeson, J., Hepburn, C., Saelen, H., 2009. Siblings, Not Triplets: Social Preferences for Risk, Inequality and Time in Discounting Climate Change. Economics: The Open-Access, Open-Assessment E-Journal 3, 1–28.

Becker, E., van den Dool, H., Zhang, Q., 2014. Predictability and Forecast Skill in NMME. Journal of Climate 27, 5891–5906.

Becker, E.J., van den Dool, H., Zhang, Q., 2013. NMME Year 2: Verification of Real-time Monthly-mean Forecasts. Science and Technology Infusion Climate Bulletin, 38th NOAA Annual Climate Diagnostics and Prediction Workshop.

van den Bijgaart, I., Gerlagh, R., Liski, M., 2016. A simple formula for the social cost of carbon. Journal of Environmental Economics and Management 77, 75–94.

Blatman, G., 2009. Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis. Ph.D. thesis. Université Blaise Pascal, Clermont-Ferrand. France.

Blatman, G., Sudret, B., 2010a. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. Probabilistic Engineering Mechanics 25, 183–197.

Blatman, G., Sudret, B., 2010b. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. Reliability Engineering & System Safety 95, 1216–1229.

Borgonovo, E., Plischke, E., 2016. Sensitivity analysis: A review of recent advances. European Journal of Operational Research 248, 869–887.

Boudoukh, J., Richardson, M., Shen, Y., Whitelaw, R.F., 2007. Do asset prices reflect fundamentals? Freshly squeezed evidence from the OJ market. Journal of Financial Economics 83, 397–412.

Brumm, J., Scheidegger, S., 2017. Using adaptive sparse grids to solve high-dimensional dynamic models. Econometrica 85, 1575–1612.

Butler, M.P., Reed, P.M., Fisher-Vanden, K., Keller, K., Wagener, T., 2014. Identifying parametric controls and dependencies in integrated assessment models using global sensitivity analysis. Environmental Modelling & Software 59, 10–29.

Cai, Y., Judd, K.L., Lontzek, T.S., 2015. The Social Cost of Carbon with Economic and Climate Risks. arXiv:1504.06909.

Calel, R., Stainforth, D.A., 2017. On the Physics of Three Integrated Assessment Models. Bulletin of the American Meteorological Society 98, 1199–1216.

Cameron, C.A., Miller, D.L., 2015. A Practitioner's Guide to Cluster-Robust Inference. Journal of Human Resources 50, 317–372.

Canova, F., 1994. Statistical Inference in Calibrated Models. Journal of Applied Econometrics 9, S123–S144.

Canova, F., 1995. Sensitivity Analysis and Model Evaluation in Simulated Dynamic General Equilibrium Economies. International Economic Review 36, 447–501.

Castelletti, A., Galelli, S., Ratto, M., Soncini-Sessa, R., Young, P., 2012. A general framework for Dynamic Emulation Modelling in environmental problems. Environmental Modelling & Software 34, 5–18.

Castruccio, S., Genton, M.G., 2016. Compressing an Ensemble With Statistical Models: An Algorithm for Global 3D Spatio-Temporal Temperature. Technometrics 58, 319–328.

Castruccio, S., McInerney, D.J., Stein, M.L., Crouch, F.L., Jacob, R.L., Moyer, E.J., 2014. Statistical emulation of climate model projections based on precomputed gcm runs. Journal of Climate 27, 1829–1844.

Chapelle, O., Vapnik, V., Bengio, Y., 2002. Model Selection for Small Sample Regression. Machine Learning 48, 9–23.

Colling, P.L., Irwin, S.H., Zulauf, C.R., 1996. Reaction of Wheat, Corn, and Soybean Futures Prices to USDA Export Inspections Reports. Review of Agricultural Economics 18, 127–136.

Collins, W.D., Bitz, C.M., Blackmon, M.L., Bonan, G.B., Bretherton, C.S., Carton, J.A., Chang, P., Doney, S.C., Hack, J.J., Henderson, T.B., Kiehl, J.T., Large, W.G., McKenna, D.S., Santer, B.D., Smith, R.D., 2006. The Community Climate System Model Version 3 (CCSM3). Journal of Climate 19, 2122–2143.

Crost, B., Traeger, C.P., 2013. Optimal climate policy: Uncertainty versus Monte Carlo. Economics Letters 120, 552–558.

DelSole, T., Tippett, M.K., 2014. Comparing Forecast Skill. Monthly Weather Review 142, 4658–4678.

Deman, G., Konakli, K., Sudret, B., Kerrou, J., Perrochet, P., Benabderrahmane, H., 2016. Using sparse polynomial chaos expansions for the global sensitivity analysis of groundwater lifetime expectancy in a multi-layered hydrogeological model. Reliability Engineering & System Safety 147, 156–169.

DeWitt, D.G., 2005. Retrospective Forecasts of Interannual Sea Surface Temperature Anomalies from 1982 to Present Using a Directly Coupled Atmosphere–Ocean General Circulation Model. Monthly Weather Review 133, 2972–2995.

Diaz, D., Moore, F., 2017. Quantifying the economic risks of climate change. Nature Climate Change 7, 774–782.

Dietz, S., Venmans, F., 2017. Cumulative carbon emissions and economic policy: in search of general principles. Grantham Research Institute WP 283.

Dringnei, D., Forest, C., Nychka, D., 2008. Parameter stimation for computationally intensive nonlinear regression with an application to climate modeling. The Annals of Applied Statistics 2, 1217–1230.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., 2004. Least angle regression. The Annals of Statistics 32, 407–499.

Egelkraut, T.M., Garcia, P., Irwin, S.H., Good, D.L., 2003. An Evaluation of Crop Forecast Accuracy for Corn and Soybeans: USDA and Private Information Agencies. Journal of Agricultural and Applied Economics 35, 79–95.

EPA, 2010. Peer Review of ADAGE and IGEM. Environmental Protection Agency, U.S., Washington D.C.

EPA, 2014. Regulatory Impact Analysis for the Proposed Carbon Pollution Guidelines for Existing Power Plants and Emission Standards for Modified and Reconstructed Power Plants. Environmental Protection Agency, U.S., Washington D.C.

EPA, 2016. Technical Support Document: Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis, August 2016. Environmental Protection Agency, U.S., Washington D.C.

Fackler, P.L., 1985. On the Relation Between Futures Price Movements and USDA Reports. Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL.

Fan, Y., van den Dool, H., 2008. A global monthly land surface air temperature analysis for 1948–present. Journal of Geophysical Research 113.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason C., Rummukainen, M., 2013. Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change: [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels,

Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Fortenbery, T.R., Sumner, D.A., 1993. The effects of USDA reports in futures and options markets. Journal of Futures Markets 13, 157–173.

Fritsch, J.M., Hilliker, J., Ross, J., 2000. Model Consensus. Weather and Forecasting 15, 571–582.

Garbade, K.D., Silber, W.L., 1983. Price Movements and Price Discovery in Futures and Cash Markets. The Review of Economics and Statistics 65, 289–297.

Ghanem, R., Higdon, D., Owhadi, H., 2017. Handbook of Uncertainty Quantification. Springer.

Golosov, M., Hassler, J., Krusell, P., Tsyvinski, A., 2014. Optimal Taxes on Fossil Fuel in General Equilibrium. Econometrica 82, 41–88.

Good, D.L., Irwin, S.H., 2006. Understanding USDA Corn and Soybean Production Forecasts: Methods, Performance and Market Impacts over 1970-2005.

Greenwald, B.C., 1983. A general analysis of bias in the estimated standard errors of least squares coefficients. Journal of Econometrics 22, 323–338.

Guinness, J., Hammerling, D., 2017. Compression and Conditional Emulation of Climate Model Output. Journal of the American Statistical Association , forthcoming.

Hagedorn, R., Doblas-Reyes, F.J., Palmer, T.N., 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting - I. Basic concept. Tellus A 57, 219–233.

Hansen, J., Sato, M., Ruedy, R., 2012. Perception of climate change. Proceedings of the National Academy of Sciences of the United States of America 109, E2415–23.

Harenberg, D., Marelli, S., Sudret, B., Winschel, V., 2019. Uncertainty Quantification and Global Sensitivity Analysis for Economic Models. Quantitative Economics, forthcoming.

Hausman, J.A., 1978. Specification Tests in Econometrics. Econometrica 46, 1251.

Hawkins, E., Smith, R.S., Gregory, J.M., Stainforth, D.A., 2016. Irreducible uncertainty in near-term climate projections. Climate Dynamics 46, 3807–3819.

Heal, G., Millner, A., 2014. Reflections: Uncertainty and Decision Making in Climate Change Economics. Review of Environmental Economics and Policy 8, 120–137.

Hicks, D.A., Thomison, P., 2004. Corn management, in: Corn: origin, history, technology, and production. John Wiley & Sons, pp. 481–522.

Holden, P., Edwards, N., 2010. Dimensionally reduced emulation of an AOGCM for application to integrated assessment modelling. Geophysical Research Letters 37, L21707.

Hope, C., 2013. Critical issues for the calculation of the social cost of CO_2: why the estimates from PAGE09 are higher than those from PAGE2002. Climatic Change 117, 531–543.

Huang, J., van den Dool, H.M., 1993. Monthly precipitation-temperature relations and temperature prediction over the united states. Journal of Climate 6, 1111–1132.

IPCC-DDC, 1998. IPCC Data Distribution Centre.

IPCC-TGICA, 2007. *General Guidelines on the Use of Scenario Data for Climate Impact and Adaptation Assessment*. Version 2. Prepared by T.R. Carter on behalf of the Intergovernmental Panel on Climate Change, Task Group on Data and Scenario Support for Impact and Climate Assessment, 66 pp.

Isengildina-Massa, O., Irwin, S.H., Good, D.L., Gomez, J.K., 2008. The Impact of Situation and Outlook Information in Corn and Soybean Futures Markets: Evidence from WASDE Reports. Journal of Agricultural and Applied Economics 40, 89–103.

Jensen, S., Traeger, C.P., 2014. Optimal climate change mitigation under long-term growth uncertainty: Stochastic integrated assessment and analytic findings. European Economic Review 69, 104–125.

Judd, K.L., Maliar, L., Maliar, S., 2011. Numerically stable and accurate stochastic simulation approaches for solving dynamic economic models. Quantitative Economics 2, 173–210.

Kann, A., Weynant, J.P., 2000. Approaches for performing uncertainty analysis in large-scale energy/economic policy models. Environmental Modeling and Assessment 5, 29–46.

Kaufmann, R.K., Kauppi, H., Mann, M.L., Stock, J.H., 2013. Does temperature contain a stochastic trend: linking statistical results to physical mechanisms. Climatic change 118, 729–743.

Kennedy, M., Kopp, S., 2000. Understanding map projections. ESRI.

Kharin, V.V., Zwiers, F.W., 2002. Climate Predictions with Multimodel Ensembles. Journal of Climate 15, 793–799.

Kirtman, B.P., Fan, Y., Schneider, E.K., 2002. The COLA Global Coupled and Anomaly Coupled Ocean–Atmosphere GCM. Journal of Climate 15, 2301–2320.

Kirtman, B.P., Min, D., Infanti, J.M., Kinter, J.L., Paolino, D.A., Zhang, Q., van den Dool, H., Saha, S., Mendez, M.P., Becker, E., Peng, P., Tripp, P., Huang, J., DeWitt,

D.G., Tippett, M.K., Barnston, A.G., Li, S., Rosati, A., Schubert, S.D., Rienecker, M., Suarez, M., Li, Z.E., Marshak, J., Lim, Y.K., Tribbia, J., Pegion, K., Merryfield, W.J., Denis, B., Wood, E.F., 2014. The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. Bulletin of the American Meteorological Society 95, 585–601.

Kloek, T., 1981. OLS Estimation in a Model Where a Microvariable is Explained by Aggregates and Contemporaneous Disturbances are Equicorrelated. Econometrica 49, 205–207.

Krakauer, N.Y., 2017. Temperature trends and prediction skill in NMME seasonal forecasts. Climate Dynamics 4, 19.

Krishnamurti, T.N., Kishtawal, C.M., LaRow, T., Bachiochi, D., Zhang, Z., Williford, E., Gadgil, S., Surendran, S., 1999. Improved Weather and Seasonal Climate Forecasts from Multimodel Superensemble. Science 285, 1548–1550.

Le Gratiet, L., Marelli, S., Sudret, B., 2017. Metamodel-based sensitivity analysis: polynomial chaos expansions and gaussian processes. handbook of uncertainty quantification, in: Handbook of Uncertainty Quantification. Springer International Publishing, pp. 1289–1325.

Leamer, E.E., 1985. Sensitivity Analyses Would Help. The American Economic Review 75, 308–313.

Lehecka, G.V., 2014. The Value of USDA Crop Progress and Condition Information: Reactions of Corn and Soybean Futures Markets. Journal of Agricultural and Resource Economics 39, 88–105.

Lobell, D.B., Roberts, M.J., Schlenker, W., Braun, N., Little, B.B., Rejesus, R.M., Hammer, G.L., 2014. Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest. Science (New York, N.Y.) 344, 513–516.

Lontzek, T.S., Cai, Y., Judd, K.L., Lenton, T.M., 2015. Stochastic integrated assessment of climate tipping points indicates the need for strict climate policy. Nature Climate Change 5, 441–444.

Lukac, L.P., Brorsen, B.W., Irwin, S.H., 1988. A test of futures market disequilibrium using twelve different technical trading systems. Applied Economics 20, 623–639.

Maier, D.E., 2004. On-farm handling, storage and quality, in: Corn: origin, history, technology, and production. John Wiley & Sons, pp. 827–866.

Maliar, L., Maliar, S., 2015. Merging simulation and projection approaches to solve high-

dimensional problems with an application to a new keynesian model. Quantitative Economics 6, 1–47.

Marelli, S., Sudret, B., 2014. UQLab: A framework for uncertainty quantification in Matlab, in: Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management, pp. 2554–2563.

Marsh, D.R., Mills, M.J., Kinnison, D.E., Lamarque, J.F., Calvo, N., Polvani, L.M., 2013. Climate Change from 1850 to 2005 Simulated in CESM1(WACCM). Journal of Climate 26, 7372–7391.

Mastrandrea, M.D., Schneider, S.H., 2004. Probabilistic Integrated Assessment of "Dangerous" Climate Change. Science 304, 571–575.

Matthews, H.D., Gillett, N.P., Stott, P.A., Zickfeld, K., 2009. The proportionality of global warming to cumulative carbon emissions. Nature 459, 829–832.

Mckay, M.D., Beckman, R.J., Conover, W.J., 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code. Technometrics 42, 55–61.

McKenzie, A.M., 2008. Pre-Harvest Price Expectations for Corn: The Information Content of USDA Reports and New Crop Futures. American Journal of Agricultural Economics 90, 351–366.

McKenzie, A.M., Holt, M.T., 2002. Market efficiency in agricultural futures markets. Applied Economics 34, 1519–1532.

McNew, K.P., Epinosa, J.A., 1994. The informational content of USDA crop reports: Impacts on uncertainty and expectations in grain futures markets. The Journal of Futures Markets 14, 475–492.

Meehl, G.A., Washington, W.M., Arblaster, J.M., Hu, A., Teng, H., Kay, J.E., Gettelman, A., Lawrence, D.M., Sanderson, B.M., Strand, W.G., 2013. Climate Change Projections in CESM1(CAM5) Compared to CCSM4. Journal of Climate 26, 6287–6308.

Meinshausen, M., Raper, S., Wigley, T., 2011a. Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6–Part 1: Model description and calibration. Atmospheric Chemistry and Physics 11, 1417–1456.

Meinshausen, M., Wigley, T., Raper, S., 2011b. Emulating atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6–Part 2: Applications. Atmospheric Chemistry and Physics 11, 1457–1471.

Merryfield, W.J., Lee, W.S., Boer, G.J., Kharin, V.V., Scinocca, J.F., Flato, G.M.,

Ajayamohan, R.S., Fyfe, J.C., Tang, Y., Polavarapu, S., 2013. The Canadian Seasonal to Interannual Prediction System. Part I: Models and Initialization. Monthly Weather Review 141, 2910–2945.

Millner, A., Calel, R., Stainforth, D.A., MacKerron, G., 2013. Do probabilistic expert elicitations capture scientists' uncertainty about climate change? Climatic Change 116, 427–436.

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P., Carter, T.R., Emori, S., Kainuma, M., Kram, T., Meehl, G.A., Mitchell, J.F.B., Nakicenovic, N., Riahi, K., Smith, S.J., Stouffer, R.J., Thomson, A.M., Weyant, J.P., Wilbanks, T.J., 2010. The next generation of scenarios for climate change research and assessment. Nature 463, 747–756.

Moulton, B.R., 1986. Random group effects and the precision of regression estimates. Journal of Econometrics 32, 385–397.

Myhre, G., Highwood, E.J., Shine, K.P., Stordal, F., 1998. New estimates of radiative forcing due to well mixed greenhouse gases. Geophysical research letters 25, 2715–2718.

National Research Council, 2010. Assessment of Intraseasonal to Interannual Climate Prediction and Predictability. The National Academies Press, Washington, DC.

Newbold, S.C., Griffiths, C., Moore, C., Wolverton, A., Kopits, E., 2013. A rapid assessment model for understanding the social cost of carbon. Climate Change Economics 04, 1350001.

Nordhaus, W.D., 1993. Rolling the DICE: an optimal transition path for controlling greenhouse gases. Resource and Energy Economics 15, 27–50.

Nordhaus, W.D., 2008. A Question of Balance: Weighing the Options on Global Warming Policies. Yale University Press.

Nordhaus, W.D., Boyer, J., 2000. Warming the world: Economic models of global warming. MIT Press, Cambridge Mass.

O'Neill, B.C., Kriegler, E., Riahi, K., Ebi, K.L., Hallegatte, S., Carter, T.R., Mathur, R., van Vuuren, D.P., 2014. A new scenario framework for climate change research: the concept of shared socioeconomic pathways. Climatic Change 122, 387–400.

Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., Church, J.A., Clarke, L., Dahe, Q., Dasgupta, P., Dubash, N.K., Edenhofer, O., Elgizouli, I., Field, C.B., Forster, P., Friedlingstein, P., Fuglestvedt, J., Gomez-Echeverri, L., Hallegatte, S., Hegerl, G., Howden, M., Jiang, K., Cisneroz, B.J., Kattsov, V., Lee, H., Mach, K.J., Marotzke, J., Mastrandrea, M.D., Meyer, L., Minx, J., Mulugetta,

Y., O'Brien, K., Oppenheimer, M., Pereira, J.J., Pichs-Madruga, R., Plattner, G.K., Pörtner, H.O., Power, S.B., Preston, B., Ravindranath, N.H., Reisinger, A., Riahi, K., Rusticucci, M., Scholes, R., Seyboth, K., Sokona, Y., Stavins, R., Stocker, T.F., Tschakert, P., van Vuuren, D., van Ypserle, J.P., 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. R. Pachauri and L. Meyer (editors), Geneva, Switzerland, IPCC. p. 151.

Palmer, T.N., Shutts, G.J., Hagedorn, R., Doblas-Reyes, F.J., Jung, T., Leutbecher, M., 2005. Representing model uncertainty in weather and climate prediction. Annual Review of Earth and Planetary Sciences 33, 163–193.

Pavan, V., Doblas-Reyes, F.J., 2000. Multi-model seasonal hindcasts over the Euro-Atlantic: skill scores and dynamic features. Climate Dynamics 16, 611–625.

Peng, P., Kumar, A., van den Dool, H., Barnston, A.G., 2002. An analysis of multimodel ensemble predictions for seasonal climate anomalies. Journal of Geophysical Research: Atmospheres 107, ACL 18–1–ACL 18–12.

Peterson, S., 2006. Uncertainty and economic analysis of climate change: A survey of approaches and findings. Environmental Modeling & Assessment 11, 1–17.

Pindyck, R.S., 2013. Climate Change Policy: What Do the Models Tell Us? Journal of Economic Literature 51, 860–872.

Pretis, F., 2015. Econometric models of climate systems: the equivalence of two-component energy balance models and cointegrated vars. University of Oxford Economics Discussion Paper 750.

Rajagopalan, B., Lall, U., Zebiak, S.E., 2002. Categorical Climate Forecasts through Regularization and Optimal Combination of Multiple GCM Ensembles. Monthly Weather Review 130, 1792–1811.

Rezai, A., Van der Ploeg, F., 2016. Intergenerational inequality aversion, growth, and the role of damages: Occam's rule for the global carbon tax. Journal of the Association of Environmental and Resource Economists 3, 493–522.

Riahi, K., van Vuuren, D.P., Kriegler, E., Edmonds, J., O'Neill, B.C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J.C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Silva, L.A.D., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J.C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., Tavoni, M., 2017.

The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. Global Environmental Change 42, 153–168.

Richardson, D.S., 2001. Measures of skill and value of ensemble prediction systems, their interrelationship and the effect of ensemble size. Quarterly Journal of the Royal Meteorological Society 127, 2473–2489.

Roll, R., 1984. Orange Juice and Weather. The American Economic Review 74, 861–880.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., Becker, E., 2014. The NCEP climate forecast system version 2. Journal of Climate 27, 2185–2208.

Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environmental Modelling & Software 25, 1508–1517.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications 181, 259–270.

Saltelli, A., D'Hombres, B., 2010. Sensitivity analysis didn't help. A practitioner's critique of the review. Global Environmental Change 20, 298–302.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global sensitivity analysis: the primer. John Wiley & Sons.

Saltelli, A., Tarantola, S., 2002. On the Relative Importance of Input Factors in Mathematical Models. Journal of the American Statistical Association 97, 702–709.

Schlenker, W., Roberts, M.J., 2006. Nonlinear Effects of Weather on Corn Yields. Review of Agricultural Economics 28, 391–398.

Slater, L.J., Villarini, G., Bradley, A.A., 2016. Evaluation of the skill of North-American Multi-Model Ensemble (NMME) Global Climate Models in predicting average and extreme precipitation and temperature over the continental USA. Climate Dynamics 1, 1–16.

Sobol, I.M., 1993. Sensitivity Estimates for Nonlinear Mathematical Models. Mathematical modelling and computational experiments 1, 407–414.

Stern, N., Peters, S., Bakhshi, V., Bowen, A., Cameron, C., Catovsky, S., Crane, D., Cruickshank, S., Dietz, S., Edmonson, N., 2006. Stern Review: The economics of climate change. HM treasury London.

Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering & System Safety 93, 964–979.

Sumner, D.A., Mueller, R.A.E., 1989. Are Harvest Forecasts News? USDA Announcements and Futures Market Reactions. American Journal of Agricultural Economics 71, 1–8.

Taylor, K.E., Stouffer, R.J., Meehl, G.A., 2012. An Overview of CMIP5 and the Experiment Design. Bulletin of the American Meteorological Society 93, 485–498.

Tebaldi, C., Knutti, R., 2007a. The use of the multi-model ensemble in probabilistic climate projections. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 365, 2053–2075.

Tebaldi, C., Knutti, R., 2007b. The use of the multi-model ensemble in probabilistic climate projections. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences 365, 2053–2075.

Thompsom, P.D., 1976. How to improve accuracy by combining independent forecasts. Monthly Weathe Review 105, 228–229.

Tol, R.S.J., 1995. The damage costs of climate change toward more comprehensive calculations. Environmental & Resource Economics 5, 353–374.

Trefethen, L.N., 2013. Approximation Theory and Approximation Practice. SIAM.

van der Ploeg, F., Rezai, A., 2019. Simple Rules for Climate Policy and Integrated Assessment. Environmental and Resource Economics 72, 77–108.

Vecchi, G.A., Delworth, T., Gudgel, R., Kapnick, S., Rosati, A., Wittenberg, A.T., Zeng, F., Anderson, W., Balaji, V., Dixon, K., Jia, L., Kim, H.S., Krishnamurthy, L., Msadek, R., Stern, W.F., Underwood, S.D., Villarini, G., Yang, X., Zhang, S., 2014. On the Seasonal Forecasting of Regional Tropical Cyclone Activity. Journal of Climate 27, 7994–8016.

Vernieres, G., Rienecker, M.M., Kovach, R., Keppenne, C.L., 2012. The GEOS-iODAS: Description and evaluation. NASA Technical Report Series on Global Modeling and Data Assimilation, NASA/ TM–2012–104606 30.

van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, Steven J an Rose, S.K., 2011. The representative concentration pathways: an overview. Climatic Change 109, 5–31.

Webster, M., 2003. Communicating Climate Change Uncertainty to Policy-Makers and the Public. Climatic Change 61, 1–8.

Webster, M., Santen, N., Parpas, P., 2012. An approximate dynamic programming frame-

work for modeling global climate policy under decision-dependent uncertainty. Computational Management Science 9, 339–362.

Weigel, A.P., Liniger, M.A., Appenzeller, C., 2008. Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? Quarterly Journal of the Royal Meteorological Society 134, 241–260.

Wesselink, A., Challinor, A.J., Watson, J., Beven, K., Allen, I., Hanlon, H., Lopez, A., Lorenz, S., Otto, F., Morse, A., Rye, C., Saux-Picard, S., Stainforth, D., Suckling, E., 2015. Equipped to deal with uncertainty in climate and impacts predictions: Lessons from internal peer review. Climatic Change 132, 1–14.

Westgate, M.E., Otegui, M.E., Andrade, F.H., 2004. Physiology of the corn plant, in: Corn: origin, history, technology, and production. John Wiley & Sons, pp. 235–271.

Wisner, R.N., Baldwin, E.D., 2004. Corn marketing, in: Corn: origin, history, technology, and production. John Wiley & Sons, pp. 753–797.

Woods, D.C., Lewis, S.M., 2017. Design of experiments for screening, in: Handbook of Uncertainty Quantification. Springer International Publishing, pp. 1143–1185.

Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT press.

Xiu, D., Karniadakis, G.E., 2002. The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations. SIAM Journal on Scientific Computing 24, 619–644.

Yeager, S.G., Shields, C.A., Large, W.G., Hack, J.J., 2006. The Low-Resolution CCSM3. Journal of Climate 19, 2545–2566.

Younes, A., Mara, T.A., Fajraoui, N., Lehmann, F., Belfort, B., Beydoun, H., 2013. Use of Global Sensitivity Analysis to Help Assess Unsaturated Soil Hydraulic Parameters. Vadose Zone Journal 12, 1–12.

Young, P.C., Ratto, M., 2011. Statistical emulation of large linear dynamic models. Technometrics 53, 29–43.

Zhang, Q., van den Dool, H., Saha, S., Mendez, M.P., Becker, E., Peng, P., Huang, J., 2011. Preliminary Evaluation of Multi-Model Ensemble System for Monthly and Seasonal Prediction. Science and Technology Infusion Climate Bulletin.

Zickfeld, K., Eby, M., Matthews, H.D., Weaver, A.J., 2009. Setting cumulative emissions targets to reduce the risk of dangerous climate change. Proceedings of the National Academy of Sciences 106, 16129–16134.

Zulauf, C.R., Irwin, S.H., 1998. Market Efficiency and Marketing to Enhance Income of Crop Producers. Review of Agricultural Economics 20, 308–331.

Curriculum Vitae of Alena Miftakhova

Personal Information

Citizenship	Russia
Birth date	21.04.1987

Education

08/2014 – 07/2019	**Ph.D. Student** in Business Administration University of Zurich, Switzerland
09/2011 – 10/2014	**Master of Arts** in Management and Economics University of Zurich, Switzerland
09/2004 – 05/2010	**Specialist** in Economics Saint-Petersburg State University of Engineering and Economics, Russia
09/2004 – 06/2009	**Bachelor of Arts** in Economics Saint-Petersburg State University of Engineering and Economics, Russia